

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/57, 9/64, C07K 16/40, C12N 1/19, 5/10, C12Q 1/68, 1/37, A61K 38/55, 39/395, G01N 33/573, 33/68		A2	(11) International Publication Number: WO 99/05291 (43) International Publication Date: 4 February 1999 (04.02.99)
(21) International Application Number: PCT/US98/15438 (22) International Filing Date: 24 July 1998 (24.07.98) (30) Priority Data: 60/053,850 25 July 1997 (25.07.97) US 60/055,836 15 August 1997 (15.08.97) US 60/062,169 16 October 1997 (16.10.97) US (71) Applicant: DU PONT PHARMACEUTICALS COMPANY [US/US]; 974 Centre Road, WR-1st 18, Wilmington, DE 19807 (US).		(72) Inventors: ARNER, Elizabeth, C. ; 386 Kelton Jennersville Road, West Grove, PA 19390 (US). BURN, Timothy, C. ; 212 Marian Court, Hockessin, DE 19707 (US). COPELAND, Robert, A. ; 20 Staten Drive, Hockessin, DE 19707 (US). DECICCO, Carl, P. ; 17 Ridgewood Turn Street, Newark, DE 19711 (US). LIU, Ruiqin ; 30 Westwood Boulevard, Hockessin, DE 19707 (US). MAGOLDA, Ronald ; 3 Church Road, Wallingford, PA 19086 (US). PRATTA, Michael ; 2 Euclid Avenue, Glassboro, NJ 08028 (US). SOLOMON, Kimberly, A. ; 437 Buttonwood Road, Landenburg, PA 19350 (US). TORTORELLA, Micky, D. ; 58 General Maxwell Court, Newark, DE 19702 (US). TRZASKOS, James, M. ; 3666 Marian Drive, Boothwyn, PA 19061 (US). YANG, Fude ; 7 Sara Court, Wilmington, DE 19810 (US). (74) Agent: KONRAD, Karen, H. ; E.I. du Pont de Nemours and Company, Legal Patent Records Center, 1007 Market Street, Wilmington, DE 19898 (US). (81) Designated States: AU, CA, IL, JP, MX, NZ, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>Without international search report and to be republished upon receipt of that report.</i>	
(54) Title: AGGRECAN DEGRADING METALLO PROTEASES			
(57) Abstract The invention is directed to the family of aggrecan degrading metallo proteases (ADMPs) that exhibit the ability to cleave the aggrecan core protein between amino acid residues Glu ³⁷³ -Ala ³⁷⁴ . The invention encompasses the nucleic acids encoding such enzymes, processes for production of recombinant ADMPs, compositions containing such enzymes, and the use of these enzymes in various assays and for the development of novel inhibitors for use as therapies for diseases involving aggrecanase-mediated degradation of cartilage or other aggrecanase-associated diseases.			

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakhstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TITLE

AGGRECAN DEGRADING METALLO PROTEASES

CROSS REFERENCE TO EARLIER FILED APPLICATION

This application claims the benefit of U.S. Provisional Application No. 60/053850 filed on July 25, 1997 and U.S. Provisional Application No. 60/055836 filed on August 15, 1997 and U.S. Provisional Application No. 60/062,169 (unknown at filing), filed on October 16, 1997.

FIELD OF THE INVENTION

The invention is directed to the family of proteins that exhibit aggrecanase activity, the nucleic acids encoding such enzymes, processes for production of recombinant aggrecanases, compositions containing such enzymes, antibodies raised against these enzymes, and the use of these enzymes or antibodies in various assays and therapies.

BACKGROUND OF THE INVENTION

Aggrecan is the major proteoglycan of cartilage and provides this tissue with its mechanical properties of compressibility and elasticity. In arthritic conditions one of the earliest changes observed in cartilage morphology is the depletion of aggrecan [Mankin et al. (1970) J. Bone Joint Surg. 52A, 424-434], which appears to be due to an increased rate of degradation.

The aggrecan molecule is composed of two N-terminal globular domains, G1 and G2, which are separated by an approximately 150 residue interglobular domain (IGD),

followed by a long central glycosaminoglycan (GAG) attachment region and a C-terminal globular domain, G3 [Hardingham et al. (1992) in Articular Cartilage and Osteoarthritis: Aggrecan, The Chondroitin Sulfate/Keratin Sulfate Proteoglycan from Cartilage (Kuettner et al.) pp. 5-20, Raven Press, New York and Paulson et al. (1987) Biochem. J. 245, 763-772]. These aggrecan molecules interact through the G1 domain with hyaluronic acid and a link protein to form large molecular weight aggregates which are trapped within the cartilage matrix [Hardingham et al. (1972) Biochim. Biophys. Acta 279, 401-405, Heinegard et al. (1974) J. Biol. Chem. 249, 4250-4256, and Hardingham, T.E. (1979) Biochem. J. 177, 237-247]. Loss of aggrecan from cartilage in arthritic conditions involves proteolytic cleavage of the aggrecan core protein within the IGD, producing a N-terminal G-1 fragment that remains bound to hyaluronic acid and the link protein within the matrix, releasing a large C-terminal GAG-containing aggrecan fragment that diffuses out of the cartilage matrix. Loss of the C-terminal fragment results in cartilage deficient in its mechanical properties. This deficiency arises because the GAGs are the components of aggrecan that impart the mechanical properties to the molecule through their high negative charge and water binding capacity.

Two major sites of proteolytic cleavage have been identified within the IGD, one between amino acid residues Asn³⁴¹-Phe³⁴² and the other between amino acid residues Glu³⁷³-Ala³⁷⁴. Although G1 fragments formed by cleavage at the Asn³⁴¹-Phe³⁴² site and at the Glu³⁷³-Ala³⁷⁴ site have been identified within articular cartilage [Flannery et al. (1992) J. Biol. Chem. 267, 1008-1014], the N-terminus identified on the large GAG-containing aggrecan C-terminal fragments in synovial fluids of patients with osteoarthritis [Sandy et al. (1992) J. Clin. Invest. 69, 1512-1516], inflammatory joint disease [Lohmander et al. (1993) Arthritis Rheum. 36, 1214-1222] and in the media from cartilage explant and chondrocyte cultures stimulated with interleukin-1 or retinoic acid [Sandy et al. (1991) J.

Biol. Chem. 266, 8198., Sandy et al. (1991) J. Biol. Chem. 266, 8683-8685., Loulakis et al. (1992) Biochem. J. 264, 589-593., Ilic et al. (1992) Arch. Biochem. Biophys. 294, 115-122., Lark et al. (1995) J. Biol. Chem. 270, 2550-2556.] was ARGSVIL, indicating that they were formed by cleavage between amino acid residues Glu³⁷³-Ala³⁷⁴. These observations suggest that cleavage at this site may be responsible for cartilage degradation.

Although many matrix metalloproteases (MMP-1, -2, -3, -7, -8, -9 and 13) have been shown to cleave in vitro at the Asn³⁴¹-Phe³⁴² site, digestion of aggrecan with a number of these purified proteases has not resulted in cleavage at the Glu³⁷³-Ala³⁷⁴ site [Fosang et al. (1992) J. Biol. Chem. 267, 19470-19474., Flannery et al. (1992) J. Biol. Chem. 267, 1008-1014., Fosang et al. (1993) Biochem. J. 295, 273-276., Fosang et al. (1996) FEBS Lett. 380, 17-20., Flannery et al. (1993) Orthop. Trans. 17, 677., and Fosang et al. (1994) Biochem. J. 305, 347-351]. Therefore, cleavage at this site has been attributed to a novel, proteolytic activity, "aggrecanase".

In addition to the Glu³⁷³-Ala³⁷⁴ bond within the interglobular domain of aggrecan, four potential aggrecanase-sensitive sites have been identified within the C-terminus of the aggrecan core protein [Loulakis et al. (1992) Biochem. J. 264, 589-593. and Sandy et al. (1995) Acta Orthop Scand (Suppl 266) 66, 26-32]. Although cleavage at these sites which are not within the interglobular domain would not be expected to release the major portion of the aggrecan molecule from the matrix, they may be involved in earlier processing of aggrecan within the matrix.

SUMMARY OF THE INVENTION

The invention encompasses a novel family of biologically active aggrecan degrading metallo proteases ("ADMP") capable of cleaving the aggrecan monomer core

protein at the Glu³⁷³-Ala³⁷⁴ aggrecanase site, as isolated and purified polypeptides. An object of the invention covers novel sequences of nucleic acids which encode for members of the ADMP family, and to expression vectors containing cDNA which encodes for novel members of the ADMP family. Another object of the invention is host cells that have been transfected or transformed with expression vectors which contain cDNA that encodes for the ADMP family of polypeptides, and processes for producing members of the ADMP family by culturing such host cells under conditions conducive to expression of an ADMP. Another object of the invention is probes containing nucleic acid sequences that hybridize to a native ADMP nucleotide sequence and the use of these probes for detection of message for an ADMP in biological samples. A further object of the invention is antibodies raised against an ADMP, which may be created as a result of the purification and isolation of members of the ADMP family and the use of such antibodies for the detection of ADMPs in biological samples. Assays utilizing an ADMP to screen for its potential inhibitors are another object of this invention. Members of the ADMP family used to design novel inhibitors of proteases exhibiting aggrecanase activity are also part of the instant invention.

Members of the ADMP family are capable of cleaving the aggrecan monomer core protein at the Glu³⁷³-Ala³⁷⁴ site, but do not readily cleave aggrecan at the Asn³⁴¹-Phe³⁴², MMP-sensitive cleavage site, and the zymogen form of the protein consists of the following domains: a propeptide domain containing a furin site, a metalloprotease domain, a disintegrin-like domain and a thrombospondin homologous domain.

As used herein, the term "zymogen" refers to the latent, full-length protein synthesized by the cells and further processed to a catalytically active form, the term "propeptide domain" refers to the N-terminal region of the molecule which contains a cysteine residue

involved in latency of the protein, the term "furin cleavage site" refers to a region of the molecule containing a tetra basic sequence of amino acids susceptible to cleavage by furin or furin-like proteases, the term "metalloprotease domain" refers to a region of the molecule which contains a zinc-binding motif with the consensus sequence, HExxHxxGxxH, responsible for the catalytic activity of the protein, the term "disintegrin-like domain" refers to a region of the molecule which exhibits sequence similarity to the disintegrin family of anti-coagulant peptides found in snake venoms, which are characterized by a high cysteine content and have the ability to disrupt cell-matrix interaction, and the term "thrombospondin homologous domain" refers to a region of the molecule containing one or more thrombospondin type 1 (TSP1) motifs with sequence homologous to the amino acid sequence of TSP1 repeats which are conserved in thrombospondin 1 and 2 and have been implicated in the interaction of thrombospondin with sulfated glycoconjugates such as heparin and heparan sulfate.

The first isolated and purified ADMP family member according to the invention, referred to as "ADMP-1", has a molecular weight between about 50 kD and about 98 kD as determined by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). More specifically, the isolated active ADMP-1 was found to have a molecular weight of approximately 67 kD as determined by SDS-PAGE. The isolated and purified metalloprotease of the invention is capable of cleaving the aggrecan monomer core protein at the Glu³⁷³-Ala³⁷⁴ site, but does not readily cleave aggrecan at the Asn³⁴¹-Phe³⁴², MMP cleavage site and consists of the following domains: a propeptide domain containing a furin site, a metalloprotease domain, a disintegrin-like domain and a thrombospondin homologous domain. The cDNA sequence of ADMP-1 is shown in SEQ ID NO:1. The isolated and purified ADMP-1 zymogen constitutes amino acids 1-837 of SEQ ID NO:2 and has a molecular weight of about 98 kDa as determined by SDS-PAGE.

The second isolated and purified aggrecanase according to the invention, referred to as "ADMP-2", has a molecular weight between about 45 kD and about 93 kD as determined by SDS-PAGE. More specifically the isolated active ADMP-2 was found as four forms of the same gene product having molecular weights of approximately 50 kD, 54 kD, 62 kD and 64 kD as determined by SDS-PAGE. This second isolated and purified metalloprotease of the invention is capable of cleaving the aggrecan monomer core protein at the Glu³⁷³-Ala³⁷⁴ site, but does not readily cleave aggrecan at the Asn³⁴¹-Phe³⁴², MMP cleavage site. The cDNA sequence of ADMP-2 is shown in SEQ ID NO:14. The isolated and purified ADMP-2 zymogen constitutes amino acids 1-930 of SEQ ID NO:15 and has a molecular weight of about 93 kDa as determined by SDS-PAGE.

The instant invention describes a method for treating a mammal having a disease characterized by an overproduction or an upregulated production of an ADMP. This treatment involves administration of a composition containing an efficacious amount of a compound that inhibits the proteolytic activity of members of the ADMP family. These enzymes include, but are not limited to, those containing the sequence of amino acids 1-837 of SEQ ID NO:2 or the sequence of amino acids from 1-930 of SEQ ID NO:15.

The potency of compounds in inhibiting soluble, active ADMP activity in conditioned media from interleukin-1-stimulated bovine nasal cartilage, correlates with their potency in inhibiting cartilage aggrecan cleavage and release from cartilage. This ADMP activity is monitored using an enzymatic assay employing purified aggrecan substrate. Specific products of aggrecanase-mediated cleavage are detected by Western analysis using the monoclonal neoepitope antibody, BC-3 [Hughes et al., Biochem. J. 306:799-804 (1995)]. This antibody recognizes the newly-formed amino-terminal sequence NH₂-ARGSVIL on fragments produced by cleavage at

the Glu³⁷³-Ala³⁷⁴ aggrecanase site. The term "neoepitope antibody" refers to an antibody which specifically recognizes a new N-terminal amino acid sequence or new C-terminal amino acid sequence generated by proteolytic cleavage but does not recognize these same sequences of amino acids when they are present within the intact protein.

Aggrecanase-mediated degradation of cartilage aggrecan has been implicated in osteoarthritis, joint injury, reactive arthritis, acute pyrophosphate arthritis (pseudogout), psoriatic arthritis and juvenile rheumatoid arthritis. Inhibitors of ADMPs would prevent cleavage of the aggrecan core protein, thereby decreasing the loss of aggrecan from the cartilage. The instant invention contains such an embodiment and also describes a method of inhibiting the cleavage of aggrecan in cartilage of a mammal by administering an efficacious amount of a compound that inhibits the aggrecanase proteolytic activity of an enzyme of the ADMP family. These enzymes include, but are not limited to, those containing the sequence of amino acids from 1-837 of SEQ ID NO:2 or the sequence of amino acids from 1-930 of SEQ ID NO:15.

Inhibitors of members of the ADMP family would be of significant clinical utility and could be potential therapeutic agents for treating the aggrecanase-related disorders cited above. ADMP inhibitors also have clinical utility for the treatment of other conditions characterized by over-production or up-regulated production of an ADMP. Isolation and purification of ADMPs would provide a significant advancement in the treatment of aggrecanase-associated diseases and in the effort to develop inhibitors of these enzymes.

DETAILED DESCRIPTION OF THE INVENTION

Isolated cDNAs encoding human ADMPs are disclosed in SEQ ID NO:1 and SEQ ID NO:14. This discovery of cDNAs encoding human ADMPs enables construction of expression vectors comprising nucleic acid sequences encoding ADMPs, host cells transfected or transformed with the expression vectors, biologically active human ADMPs as isolated and purified proteins, and antibodies immunoreactive with ADMPs.

Isolated and purified ADMP polypeptides according to the invention are useful for detecting the aggrecanase-inhibiting activity of a molecule. In such a method involving routine and conventional techniques, a molecule of unknown aggrecanase-inhibiting activity is mixed with a substrate and incubated with an ADMP polypeptide. The extent of substrate cleavage then can be determined using a neoepitope antibody to detect cleavage fragments generated by specific cleavage at the Glu³⁷³-Ala³⁷⁴ bond.

In addition, ADMP polypeptides according to the invention are useful for the structure-based design of an aggrecanase inhibitor. Such a design would comprise the steps of determining the three-dimensional structure of such ADMP polypeptide, analyzing the three-dimensional structure for the likely binding sites of substrates, synthesizing a molecule that incorporates a predictive reactive site, and determining the aggrecanase-inhibiting activity of the molecule.

Antibodies immunoreactive with ADMPs are now made available through the invention. Such antibodies may be useful for inhibiting aggrecanase activity in vivo and for detecting the presence of an ADMP in a sample.

As used herein, the term "ADMP" refers to a family of polypeptides that are capable of cleaving the aggrecan core protein at the Glu³⁷³-Ala³⁷⁴ bond, but do not readily cleave at the Asn³⁴¹-Phe³⁴² bond and consist of the following domains: a propeptide domain containing a furin

site, followed by a metalloprotease domain, followed by a disintegrin-like domain, followed by a thrombospondin homologous domain, wherein the polypeptide is either a native or recombinant polypeptide.

The ADMP family encompasses, but is not limited to, proteins having the amino acid sequence 1 to 837 of SEQ ID NO:2 or the sequence of amino acids from 1-930 of SEQ ID NO:15, as well as those proteins having a high degree of similarity (at least 80% homology) with the amino acid sequence 1 - 837 of SEQ ID NO: 2 or the sequence of amino acids from 1-930 of SEQ ID NO:15 and which proteins are biologically active. In addition, ADMP refers to the biologically active gene products of the nucleotides 405 - 2919 of SEQ ID NO: 1 and to the biologically active gene products of the nucleotides 121-2910 of SEQ ID NO:14. Further encompassed by the term "ADMP" are the truncated proteins that retain biological activity. Truncated versions are those having less of the C-terminal or N-terminal portion of the protein or that comprise substantially all of the catalytic domain, i.e., amino acids 212-431 of SEQ NO:2 or amino acids 262-479 of SEQ ID NO:15.

The first isolated and purified active aggrecanase according to the invention, referred to as "ADMP-1", has a molecular weight between about 50 kD and about 98 kD as determined by sodium doecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). More specifically, ADMP-1 was found to have a molecular weight of approximately 67 kD as determined by SDS-PAGE. The second isolated and purified active aggrecanase according to the invention, referred to as "ADMP-2", has a molecular weight between about 45 kD and about 93 kD as determined by SDS-PAGE. More specifically, active ADMP-2 was present as four forms of the same gene product found to have molecular weights of approximately 50 kD, approximately 54 kD, approximately 62 kD and approximately 64 kD as determined by SDS-PAGE.

The term "isolated" as used herein, means that an ADMP is essentially free from association with other proteins or polypeptides, for example as a purification product of recombinant host cell culture or as a purified product from a non-recombinant source. The term "substantially purified" as used herein, refers to a mixture that contains an ADMP and is essentially free from association with other proteases, and which substantially purified ADMP retains biological activity. The term "purified ADMP" means that the ADMP is present in a cell-free system. The term "biologically active" as it refers to an ADMP, means that the ADMP is capable of cleaving the aggrecan core protein at the Glu³⁷³-Ala³⁷⁴ bond.

This invention provides a nucleic acid molecule encoding ADMP-1 (SEQ ID NO:1) and ADMP-2 (SEQ ID NO:14). Examples of nucleic acid molecules are RNA, cDNA or isolated genomic DNA molecules. One means of isolating an ADMP is to probe a cDNA or genomic library with a natural or artificial DNA probe derived from the ADMP-1 or ADMP-2 cDNA. DNA probes derived from the ADMP-1 or ADMP-2 cDNA can be used to obtain complementary cDNA, RNA or genomic clones from human, mammalian or other sources, using methods known in the art (such as those outlined in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989)).

A "nucleotide sequence" refers to a polynucleotide molecule in the form of a separate fragment or as a component of a larger nucleic acid construct, that has been derived from DNA or RNA isolated at least once in substantially pure form (i.e., free of contaminating endogenous materials and in a quantity or concentration enabling identification, manipulation, and recovery of its component nucleotide sequences by biochemical methods (such as those outlined in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989)). Such

sequences are preferably provided in the form of an open reading frame uninterrupted by internal non-translated sequences, or introns, that are typically present in eukaryotic genes. Sequences of non-translated DNA may be present 5' or 3' from an open reading frame, where the same do not interfere with manipulation or expression of the coding region.

The term "aggrecan degrading metallo protease" ("ADMP") as referred to herein, means a polypeptide substantially homologous to a native ADMP which is biologically active, and an amino acid sequence different from that of the native ADMP (human, bovine, canine, murine or other species) because of one or more deletions, insertions or substitutions. The term includes a variant sequence wherein the variant amino acid sequence preferably is at least 80% identical to a native ADMP amino acid sequence. The percent identity may be determined, for example, by comparing sequence information using the GAP computer program, version 6.0 described by Devereux et al. (*Nucl. Acids Res.* 12:387, 1984) and available from the University of Wisconsin Genetics Computer Group (UWGCG). The GAP program utilizes the alignment method of Needleman and Wunsch (*J. Mol. Biol.* 48:443, 1970), as revised by Smith and Waterman (*Adv. Appl. Math* 2:482, 1981). The preferred default parameters for the GAP program include: (1) a unary comparison matrix (containing a value of 1 for identities and 0 for non-identities) for nucleotides, and the weighted comparison matrix of Gribskov and Burgess, *Nucl. Acids Res.* 14:6745, 1986, as described by Schwartz and Dayhoff, eds., *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, pp. 353-358, 1979; (2) a penalty of 3.0 for each gap and an additional 0.10 penalty for each symbol in each gap; and (3) no penalty for end gaps.

"Variants" as referred to herein comprise conservatively substituted sequences, meaning that a given amino acid residue is replaced by a residue having

similar physiochemical characteristics. Conservative substitutions are well known in the art and include substitution of one aliphatic residue, such as Ile, Val, Leu, or Ala for another, or substitutions of one polar residue for another, such as between Lys and Arg, Glu and Asp, or Gln and Asn. Conventional procedures and methods can be used for making and using such variants. Other such conservative substitutions, for example, substitutions of entire regions having similar hydrophobicity characteristics, are well known and routinely performed. Naturally occurring ADMP variants are also encompassed by the invention. Examples of such variants are proteins that result from alternate mRNA splicing events or from proteolytic cleavage of an ADMP, wherein the aggrecanase proteolytic property is retained. Alternate splicing of mRNA may yield a truncated but biologically active ADMPs. Variations attributable to proteolysis include, for example, differences in the N- or C-termini upon expression in different types of host cells, due to proteolytic removal of one or more terminal amino acids from the ADMP.

As stated above, the invention provides isolated and purified, or homogeneous, ADMP polypeptides, both recombinant and non-recombinant. Variants and derivatives of native ADMPs that retain the desired biological activity may be obtained by mutations of nucleotide sequences coding for native ADMP polypeptides. Alterations of the native amino acid sequence may be accomplished by any of a number of conventional methods. Mutations can be introduced at particular loci by synthesizing oligonucleotides containing a mutant sequence, flanked by restriction sites enabling ligation to fragments of the native sequence. Following ligation, the resulting reconstructed sequence encodes an analog having the desired amino acid insertion, substitution, or deletion.

Alternatively, oligonucleotide-directed procedures can be employed to provide an altered gene wherein

predetermined codons can be altered by substitution, deletion or insertion. Exemplary methods of making the alterations set forth above are disclosed by Walder et al. (*Gene* 42:133, 1986); Bauer et al. (*Gene* 37:73, 1985); Craik (*BioTechniques*, January 1985, 12-19); Smith et al. (*Genetic Engineering: Principles and Methods*, Plenum Press, 1981); Kunkel (*Proc. Natl. Acad. Sci. USA* 82:488, 1985); Kunkel et al. (*Methods in Enzymol.* 154:367, 1987); and Mark et al. (*Proc. Natl. Acad. Sci. USA* 18:5662, 1984) all of which are incorporated by reference.

An ADMP may be modified to create ADMP derivatives by forming covalent or aggregative conjugates with other chemical moieties, such as glycosyl groups, polyethylene glycol (PEG) groups, lipids, phosphate, acetyl groups and the like. Covalent derivatives of an ADMP may be prepared by linking the chemical moieties to functional groups on the ADMP amino acid side chains or at the N-terminus or C-terminus of an ADMP polypeptide. Other derivatives of an ADMP within the scope of this invention include covalent or aggregative conjugates of an ADMP or its fragments with other proteins or polypeptides, such as by synthesis in recombinant culture as N-terminal or C-terminal fusions. For example, the conjugate may comprise a signal or leader polypeptide sequence (e.g. the α -factor leader of *Saccharomyces*) at the N-terminus of an ADMP polypeptide. The signal or leader peptide co-translationally or post-translationally directs transfer of the conjugate from its site of synthesis to a site inside or outside of the cell membrane or cell wall.

ADMP polypeptide conjugates can comprise peptides added to facilitate purification and identification of the ADMP. Such peptides include, for example poly-His or the antigenic identification peptides described in Hopp et al., *Bio/Technology*.6:1204, 1988. The term "ADMP derivative" refers to an ADMP polypeptide conjugated with a chemical moiety, other proteins or polypeptides encompassing, but not limited to, glycosyl groups, polyethylene glycol (PEG) groups, lipids, phosphate,

acetyl groups, poly-His peptides, antigenic-identification peptides, signal peptides or leader peptides.

The invention further includes ADMP polypeptides with or without associated native-pattern glycosylation. An ADMP expressed in yeast or mammalian expression systems (e.g., COS-7 cells) may be similar to or significantly different from the native ADMP polypeptide in molecular weight and glycosylation pattern, depending upon the choice of expression system. Expression of ADMP polypeptides in expression systems, such as *E. coli*, provides non-glycosylated molecules. Glycosyl groups may be removed through conventional methods, in particular those utilizing glycopeptidase. In general, a glycosylated ADMP may be incubated with a molar excess of glycopeptidase.

Equivalent DNA constructs that encode various additions or substitutions of amino acid residues or sequences, or deletions of terminal or internal residues or sequences not needed for biological activity, are encompassed by the invention. For example, N-glycosylation sites in the ADMP extracellular domain can be modified to preclude glycosylation, allowing expression of a reduced carbohydrate analog in mammalian and yeast expression systems. N-glycosylation sites in eukaryotic polypeptides are characterized by an amino acid triplet Asn-X-Y, wherein X is any amino acid except Pro and Y is Ser or Thr. Appropriate substitutions, additions or deletions to the nucleotide sequence encoding these triplets will result in prevention of attachment of carbohydrate residues at the Asn side chain. Alteration of a single nucleotide, chosen so that Asn is replaced by a different amino acid for example, is sufficient to inactivate the N-glycosylation site. Known procedures for inactivating N-glycosylation sites in proteins include those described in Larsen et al. (*J. Biol. Chem.* 263:1023, 1988), Hansen et al. (*J. Biol. Chem.* 263:15713, 1988) and Larsen et al. (*Blood* 73:1842, 1989), hereby incorporated by reference.

In another example, sequences encoding Cys residues that are not essential for biological activity can be altered to cause the Cys residues to be deleted or replaced with other amino acids, preventing formation of incorrect intramolecular disulfide bridges upon renaturation. Other equivalents may be prepared by modification of adjacent dibasic amino acid residues to enhance expression in yeast systems in which protease activity is present.

Nucleic acid sequences within the scope of the invention include isolated DNA and RNA sequences that hybridize to the native ADMP nucleotide sequences disclosed herein or those of other members of the ADMP family under conditions of moderate or high stringency, and which encode a biologically active ADMP. Conditions of moderate stringency, as known to those having ordinary skill in the art, and as defined by Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2 ed. Vol. 1, pp. 101-104, Cold Spring Harbor Laboratory Press, (1989), include use of a prewashing solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0) and hybridization conditions of about 50°C - 60°C, 5 X SSC, overnight, preferably 55°C. Conditions of high stringency include higher temperatures of hybridization and washing. The skilled artisan will recognize that the temperature and wash solution salt concentration may be adjusted as necessary according to factors such as the length of the probe.

Due to the known degeneracy of the genetic code wherein more than one codon can encode the same amino acid, a DNA sequence may vary from that shown in SEQ ID NO:1 and still encode an ADMP having the amino acid sequence of SEQ ID NO:2 or a DNA sequence may vary from that shown in SEQ ID NO:14 and still encode an ADMP having the amino acid sequence of SEQ ID NO:15. Such variant DNA sequences may result from silent mutations (e.g., occurring during PCR amplification), or may be the product of deliberate mutagenesis of a native sequence.

The invention thus provides equivalent isolated DNA sequences encoding biologically active ADMPs, selected from: (a) the coding region of a native ADMP gene, (b) cDNA comprising the nucleotide sequence presented in SEQ ID NO:1 or comprising the nucleotide sequence presented in SEQ ID NO:14, (c) DNA capable of hybridization to a DNA of (a) or (b) under moderately stringent conditions and which encodes a biologically active ADMP, and (d) DNA which is degenerate as a result of the genetic code to a DNA defined in (a), (b) or (c) and which encodes a biologically active ADMP. ADMPs encoded by such DNA. equivalent sequences are encompassed by the invention.

DNAs that are equivalents to the DNA sequence of SEQ ID NO:1 will hybridize under moderately stringent or highly stringent conditions to the double-stranded native DNA sequence that encode polypeptides comprising amino acid sequences of 1 to Xaa of SEQ ID NO:2, wherein Xaa is an amino acid from 431-837. DNAs that are equivalents to the DNA sequence of SEQ ID NO:14 will hybridize under moderately stringent or highly stringent conditions to the double-stranded native DNA sequence that encode polypeptides comprising amino acid sequences of 1 to Xaa of SEQ ID NO:15, wherein Xaa is an amino acid from 479-930. Examples of ADMP proteins encoded by such DNA, include, but are not limited to, ADMP fragments and ADMPs comprising inactivated N-glycosylation site(s), inactivated KEX2 protease processing site(s), or conservative amino acid substitution(s), as described above. ADMP proteins encoded by DNA derived from other species, wherein the DNA will hybridize under conditions of moderate or high stringency to the complement of the cDNA of SEQ ID NO:1 or the cDNA of SEQ ID NO:14 are also encompassed by this invention.

ADMP polypeptides may exist as oligomers, such as covalently-linked or non-covalently-linked dimers or trimers. Oligomers may be linked by disulfide bonds formed between cysteine residues on different aggrecanase polypeptides. In one embodiment of the invention, an

ADMP dimer is created by fusing the ADMP to the Fc region of an antibody (e.g., IgG1) in a manner that does not interfere with biological activity of the ADMP. The Fc polypeptide preferably is fused to the C-terminus of a soluble ADMP. General preparation of fusion proteins comprising heterologous polypeptides fused to various portions of antibody-derived polypeptides (including the Fc domain) has been described, e.g., by Ashkenazi et al. (*PNAS USA* 88:10535, 1991) and Byrn et al. (*Nature* 344:677, 1990). A gene fusion encoding the ADMP:Fc fusion protein is inserted into an appropriate expression vector. ADMP:Fc fusion proteins are allowed to assemble much like antibody molecules, whereupon interchain disulfide bonds form between Fc polypeptides, yielding a divalent ADMP. If fusion proteins are made with both heavy and light chains of an antibody, it is possible to form an aggrecanase oligomer with as many as four aggrecanase molecules. Alternatively, one can link two soluble aggrecanase domains with a peptide linker.

Expression vectors containing a nucleic acid sequence encoding an ADMP can be utilized to produce recombinant protein. An ADMP DNA sequence can be operably linked to suitable transcriptional and translational regulatory nucleotide sequences using established procedures. Regulatory sequences, which are usually derived from viral, mammalian or insect genes, can include transcriptional promoters, operators, or enhancers, mRNA ribosomal binding sites, and/or other appropriate sequences which drive transcription, translation initiation and termination. When a regulatory sequence is functionally related to the ADMP DNA sequence, the nucleotide sequence is operably linked. Thus, a promoter nucleotide sequence is operably linked to an ADMP DNA sequence if the promoter nucleotide sequence drives the transcription of the ADMP DNA sequence. The expression vector may additionally include an origin of replication, to mediate replication in the desired host cells, as well as a selectable marker gene

Expression vectors may also include signal peptide sequences (secretory leaders), which may be fused in-frame to the ADMP sequence. The inclusion of the signal sequence on the resultant fusion protein can enhance extracellular secretion of the ADMP polypeptide. The signal peptide may be cleaved from the ADMP protein upon export through the cellular secretory pathway.

Prokaryotes include gram negative or gram positive organisms, for example, *E. coli* or Bacilli. Suitable prokaryotic host cells for transformation include, for example, *E. coli*, *Bacillus subtilis*, *Salmonella typhimurium*, and various other species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*. In a prokaryotic host cell, such as *E. coli*, an ADMP polypeptide may include an N-terminal methionine residue to facilitate expression of the recombinant polypeptide in the prokaryotic host cell. The N-terminal Met may be cleaved from the expressed recombinant ADMP polypeptide.

18

vector pBR322 contains genes for ampicillin and tetracycline resistance and thus provides simple means for identifying transformed cells. To construct an expression vector using pBR322, an appropriate promoter and an ADMP DNA sequence are inserted into the pBR322 vector. Other commercially available vectors include for example, pET (Novagen, Madison, WI, USA) and PGEMI (Promega Biotec, Madison, WI, USA).

Promoter sequences commonly used for recombinant prokaryotic host cell expression vectors include B-lactamase (penicillinase), lactose promoter system (Chang et al., *Nature* 275:615, 1978; and Goeddel et al., *Nature* 281:544, 1979), a phage λ P_L promoter, tryptophan (trp) promoter system (Goeddel et al., *Nucl. Acids Res.* 8:4057, 1980) and tac promoter (Maniatis, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, p. 412, 1982).

ADMP polypeptides alternatively may be expressed in yeast host cells, preferably from the *Saccharomyces* genus (e.g., *S. cerevisiae*). Other genera of yeast, such as *Pichia K.lactis* or *Kluyveromyces*, may also be employed. Yeast vectors will often contain an origin of replication sequence from a 2 μ yeast plasmid, an autonomously replicating sequence (ARS), a promoter region, sequences for polyadenylation, sequences for transcription termination, and a selectable marker gene. Suitable promoter sequences for yeast vectors include among others, promoters for metallothionein, 3-phosphoglycerate kinase (Hitzeman et al., *J. Biol. Chem.* 255:2073, 1980) or other glycolytic enzymes (Hess et al., *J. Adv. Enzyme Reg.* 7,149, 1968; and Holland et al., *Biochem.* L7:4900, 1978), such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase. Other appropriate vectors and promoters for use in yeast expression are described in Hitzeman, EPA-73,657 or in

Fleer et al., *Gene*, 107:285-195 (1991); and van den Berg et. al., *Bio/Technology*, 8:135-139 (1990). Another alternative is the glucoserepressible ADH2 promoter described by Russell et al. (*J. Biol. Chem.* 258:2674, 1982) and Beier et al. (*Nature* 300:724, 1982). Shuttle vectors replicable in both yeast and *E. coli* may be constructed by inserting DNA sequences from pBR322 for selection and replication in *E. coli* (Amp^r gene and origin of replication) into the above-described yeast vectors.

The yeast α -factor leader sequence, typically inserted between the promoter and the cDNA to be expressed, may be employed to mediate secretion of an ADMP polypeptide. See, e.g., Kurjan et al., *Cell* 3D:933, 1982; Bitter et al., and *Proc. Natl. Acad. Sci. USA* 11:53301, 1984. Other leader sequences suitable for facilitating secretion of recombinant polypeptides from yeast hosts are known to those skilled in the art. A leader sequence may be modified near its 3' end to contain one or more restriction sites. This will facilitate fusion of the leader sequence to the structural gene.

Yeast transformation protocols are known to those skilled in the art. One such protocol is described by Hinnen et al., *Proc. Natl. Acad. Sci. USA* 11:1929, 1978. The Hinnen et al. protocol selects for Trp⁺ transformants in a selective medium, wherein the selective medium consists of 0.67% yeast nitrogen base, 0.5% casamino acids, 2% glucose, 10 ug/ml adenine and 20 ug/ml uracil.

Yeast host cells transformed by vectors containing ADH2 promoter sequence may be grown for inducing expression in a "rich" medium. An example of a rich medium is one consisting of 1% yeast extract, 2% peptone, and 1% glucose supplemented with 80 ug/ml adenine and 80 ug/ml uracil. Derepression of the ADH2 promoter occurs when glucose is exhausted from the medium.

The expression of recombinant ADMP polypeptides can also be carried out in mammalian or insect host cell

culture systems. Established cell lines of mammalian origin may also be employed. Examples of suitable mammalian host cell lines include L cells, C127 cells, 3T3 cells (ATCC CCL 163), the COS-7 line of monkey kidney cells (ATCC CRL 1651) (Gluzman et al., *Cell* 2.1:175; 1981), HeLa cells, Chinese hamster ovary (CHO) cells, and BHK (ATCC CRL 10) cell lines, and the CV-1/EBNA-1 cell line derived from the African green monkey kidney cell line CV1 (ATCC CCL 70) as described by McMahan et al. (*EMBO J.* 10: 2821, 1991). Baculovirus systems for production of heterologous proteins in insect cells are reviewed by Luckow and Summers, *Bio/Technology* 6:47 20 (1988).

Generally, the expression of eukaryotic genes in mammalian host cells is driven by viral-genome-derived early and late promoters, enhancer, splice signals and polyadenylation sites, which are included in a variety of mammalian expression vectors. Viral early and late promoters are particularly useful because both are easily obtained from a viral genome as a fragment which may also contain a viral origin of replication (Fiers et al., *Nature* 273: 113, 1978). Typically used viral transcriptional and translational control sequences are derived from Rous sarcoma virus, Polyoma virus, Adenovirus 2, Simian Virus 40 (SV40), and human cytomegalovirus (CMV).

An isolated and purified ADMP protein according to the invention may be produced by recombinant expression systems as described above or purified from media of stimulated tissue or cells. ADMPs can be substantially purified, as indicated by a single protein band upon analysis by SDS-polyacrylamide gel electrophoresis (SDS-PAGE). One process for producing an ADMP consists of culturing a host cell transformed with an expression vector containing a DNA sequence that encodes an ADMP under conditions sufficient to promote expression of the ADMP. The ADMP is then recovered from culture medium or cell extracts, depending on the expression system used.

As known to one skilled in the art, procedures for purifying a recombinant protein will vary according to such factors as the type of host cells used and whether or not the recombinant protein is secreted into the culture medium. For example, when expression systems that secrete the recombinant protein are employed, the culture medium first may be concentrated using a commercially available protein concentration filter, for example, an Amicon or Millipore, Pellicon ultrafiltration unit. Following the concentration step, the concentrate can be applied to a purification matrix such as a gel filtration medium, or an anion or a cation exchange resin of the type commonly used in protein purification. Finally, one or more reverse-phase high-performance liquid chromatography (RP-HPLC) steps utilizing hydrophobic RP-HPLC media can be employed to further purify the ADMP. Some or all of the foregoing purification steps, in various combinations, can be used to provide an isolated and purified recombinant protein.

In addition to recombinantly producing ADMPs, ADMPs may be isolated and purified from conditioned media of stimulated bovine nasal cartilage cultures, such stimulation effected with cytokines such as IL-1 or TNF, retinoic acid, adhesion molecule fragments such as fibronectin fragments or other stimuli. Other sources of aggrecanase may be used, including but not limited to, cartilage and other aggrecanase-expressing tissues from various species, and ADMPs may also be produced by stimulated cells in culture. ADMP probes containing nucleic acid sequences that hybridize to native ADMP nucleotide sequence, available through the invention, can be used to enable identification of cell lines, cells or tissue sources of ADMPs. Once a source of ADMPs is identified, ADMPs may be isolated and purified by optimally stimulating the source cells or tissue to produce ADMPs.

It is possible to utilize an affinity column comprising an ADMP-binding protein to affinity-purify

expressed ADMP polypeptides. ADMP polypeptides can be removed from an affinity column using conventional techniques, e.g., in a high salt elution buffer and then dialyzed into a lower salt buffer for use or by changing pH or other components depending on the affinity matrix utilized. Example 5 describes a procedure for employing ADMPs of the invention to generate antibodies directed against the ADMPs.

Recombinant protein produced in bacterial culture is usually isolated by initial disruption of the host cells, centrifugation, extraction from cell pellets if an insoluble polypeptide, or from the supernatant fluid if a soluble polypeptide. This isolation is followed by one or more concentrating steps such as salting-out, ion exchange, affinity purification or size exclusion chromatography. RP-HPLC can be employed for final purification steps. Microbial cells can be disrupted by any convenient method including freeze-thaw cycling, sonication, mechanical disruption, or use of cell-lysing agents.

Transformed yeast host cells are preferably used to express an ADMP as a secreted polypeptide in order to simplify purification. Secreted recombinant polypeptide from a yeast host cell fermentation can be purified by methods analogous to those disclosed by Urdal et al. (*J. Chromatog.* 2k:171, 1984).

Antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to a target ADMP mRNA sequence (forming a duplex) or to the ADMP sequence in the double-stranded DNA helix (forming a triple helix) can be made according to the invention. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of an ADMP cDNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to about 30 nucleotides. The ability to create an antisense or a sense oligonucleotide, based upon a cDNA sequence for a

given protein is described in, for example, Stein and Cohen, *Cancer Res.* 48:2659, 1988 and van der Krol et al., *BioTechniques* 6:958, 1988.

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of complexes that block translation (RNA), or transcription (DNA) by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation. Thus, the antisense oligonucleotides may be used to block expression of ADMP proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences. Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, CaPO₄-mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. Antisense or sense oligonucleotides are preferably introduced into a cell containing the target nucleic acid sequence by insertion of the antisense or sense oligonucleotide into a suitable retroviral vector, then contacting the cell with the

retrovirus vector containing the inserted sequence, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C.

Sense or antisense oligonucleotides also may be introduced into a cell containing the nucleotide sequence by formation of a conjugate with a ligand binding molecule. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

Detection of ADMP enzymatic activity in crude culture media from tissue or cell cultures or partially-purified or purified ADMP preparations can be achieved by incubating the ADMP-containing material with an aggrecan substrate and monitoring the production of aggrecan fragments generated by specific cleavage at the Glu373-Ala374 bond using a neoepitope antibody to either the new N-terminus, ARGSV, or the new C-terminus, NITEGE, formed by cleavage at this bond. The ARGSV neoepitope antibodies used encompass, but are not limited to, the BC-3 monoclonal antibody (Hughes, C.E., et al., Biochem. J. 306:799-804, 1995). As used herein, "aggrecan" refers to the aggregating proteoglycan, aggrecan, from human or animal cartilage, as the native aggrecan isolated from tissue, as recombinant full-length aggrecan or as a

recombinant protein representing a portion of the aggrecan molecule. Within an aspect of the invention, an ADMP may be utilized to identify additional ADMP-sensitive sites thus enabling activity to also be detected by monitoring the production of fragments formed by cleavage at alternative ADMP-sensitive sites using neoepitope antibodies to the new C-terminus or to the new N-terminus generated by ADMP-specific cleavage at these sites. Alternative sites in the aggrecan core protein encompass, but are not limited to the E1545-G1546, E1714-G1715, E1819-A1820, or E1919-L1920 bond (numbering based on the human aggrecan-core protein sequence). These human aggrecan ADMP-sensitive cleavage sites are conserved in aggrecan from various animal species although the absolute numbering based on the sequence of the aggrecan core protein may vary from species to species. Conserved amino acid sequences in various species around conserved ADMP-sensitive sites are shown below.

Human	NITEGE ³⁷³	ARGSVILT ³⁷⁴
Bovine	NITEGE	ARGSVILT
Rat	NITEGE	ARGNVILT
Mouse	NVTEGE	ALGSVILT
Pig	NITEGE	ARGTVILT
Sheep	NITEGE	ARGNVILT
Chicken	NVTEEE	ARGSI
Horse	NITEGE	ARGNVILT
Human	ASTASELE ¹⁵⁴⁵	GRGTIGIS ¹⁵⁴⁶
Bovine	ATTAGELE	GRGTIDIS
Mouse	ATTSSLE	GRGTIGIS
Rat	ATTASELE	GRGTISVS
Human	PTTFKEEE ¹⁷¹⁴	GLGSVELS ¹⁷¹⁵
Bovine	PTTFKEEE	GLGSVELS
Rat	PTTFREEE	GLGSVELS
Mouse	PTTFREEE	GLGSVELS
Human	TQAPTAQE ¹⁸¹⁹	AGEGPSGI ¹⁸²⁰
Bovine	TQAPTAQE	AGEGPSGI
Rat	TLAPTAQE	AGEGPSSI
Mouse	TQAPTAQE	AGEGPSGI
Chicken	TQTSVAQE	VGEGPSGM

Human	TEPTISQE ¹⁹¹⁹	¹⁹²⁰ LGQRPPVT
Bovine	TEPTVSQE	LGQRPPVT
Rat	TEPTVSQE	LGHGPSMT
Mouse	TEPTVSQE	LGHGPSMT
Chicken	TRPTVSQE	LGGETAVT
Dog	TEPTVSQE	LAQRPPVT

Thus, aggrecan from various animal species, including but not limited to, bovine, dog, pig, rat, mouse, sheep, horse and chicken may also be used as a substrate for detecting ADMP activity. Utilizing neoepitope antibodies allows detection of fragments formed specifically by ADMP-mediated cleavage even in the presence of other proteolytic activities that may be present in crude preparations.

As used herein, the cleavage site "E373-374A" refers to the ITEGE373-374ARGS bond of human aggrecan as well as to the homologous aggrecanase-sensitive cleavage site in aggrecan from various animal species, the cleavage site "E1545-1546G" refers to the SELE1545-1546GRGT bond of human aggrecan as well as to the homologous aggrecanase-sensitive cleavage site in aggrecan from various animal species, the cleavage site "E1714-1715G" refers to the KEEE1714-1715GLGS bond of human aggrecan as well as to the homologous aggrecanase-sensitive cleavage site in aggrecan from various animal species, the cleavage site "E1819-1820A" refers to the TAQE1819-1820AGEG bond of human aggrecan as well as to the homologous aggrecanase-sensitive cleavage site in aggrecan from various animal species, the cleavage site "E1919-1920L" refers to the ISQE1919-1920LGQR bond of human aggrecan as well as to the homologous aggrecanase-sensitive cleavage site in aggrecan from various animal species.

A purified ADMP may also be assayed using any of a variety of protease assays known in the art. In general, an ADMP can be assayed through the use of a peptide substrate that represents the natural cleavage site of aggrecan cleavage. For example, in order to detect the cleavage of a substrate by an ADMP, the substrate can be

tagged with a fluorescent group on one side of the cleavage site and with a fluorescence-quenching group on the opposite side of the cleavage site. Upon cleavage by the ADMP, quenching is eliminated thus providing a detectable signal. Alternatively, the substrate may be tagged with a calorimetric leaving group that more strongly absorbs upon cleavage. Alternatively, the substrate may have a thioester group synthesized into the cleavage site of the substrate so that upon cleavage by an ADMP, the thiol group remains and can be easily detected using conventional methods.

Within an aspect of the invention, an ADMP and peptides based on the amino acid sequence of the ADMP, may be utilized to prepare antibodies that specifically bind to the ADMP. Specific examples of such antibody preparation is described in Example 5 and 6 herein. The term "antibodies" is meant to include polyconal antibodies, monoclonal antibodies, fragments thereof such as F(ab')₂, and Fab, as well as any recombinantly produced binding parameters. Antibodies are defined to be specifically binding if they bind an ADMP with a K_a of greater than or equal to about 1×10^{-7} M. Affinities of binding partners or antibodies can be readily determined using conventional techniques, for example those described by Scatchard et al., *Ann. N.Y. Acad. Sci.*, 51:660 (1949).

Using standard procedures, polyclonal antibodies can be readily generated from a variety of sources such as horses, cows, goats, sheep, dogs, chickens, rabbits, mice or rats. In general a purified ADMP, or a peptide based on the amino acid sequence of the ADMP, that is appropriately conjugated is administered to the host animal typically through parenteral injection. The immunogenicity of the ADMP may be enhanced by the use of an adjuvant such as Freund's complete or incomplete adjuvant. Following booster immunizations, small samples of serum are collected and tested for reactivity to the ADMP or the ADMP peptides. Examples of various assays

useful for such determination include those described in: *Antibodies: A Laboratory Manual*, Harlow and Lane (eds.), Cold Spring Harbor Laboratory Press, 1988; as well as procedures such as countercurrent immuno-electrophoresis (CIEP), radioimmunoassay radio-immunoprecipitation, enzyme-linked immuno-sorbent assays (ELISA), dot-blot assays, and sandwich assays.

Monoclonal antibodies can be readily prepared using standard procedures such as those described in *Monoclonal Antibodies, Hybridomas: A New Dimension in Biological Analyses*, Plenum Press, Kennett, McKearn, and Bechtol (eds.), 1980. The host animals, for example mice, are injected intraperitoneally at least once, and preferably at least twice at approximate 3 week intervals, with an isolated and purified ADMP or conjugated ADMP peptide, optionally in the presence of adjuvant. Mouse sera are then assayed by the conventional dot-blot technique or antibody-capture technique, to determine which animal is best to use in the production of hybrid cells. Approximately two to three weeks later, the mice are given an intravenous boost of the ADMP or conjugated ADMP peptide. Mice are subsequently sacrificed and using established protocols, the spleen cells are fused with commercially available myeloma cells. The fusing agent can be any suitable agent used in the art, for example, polyethylene glycol. The fused cells are spread onto plates containing media that allows for their selective growth. The fused cells are grown for approximately eight days. Supernatant from the resulting hybridomas is collected and added to a plate that has been coated with goat anti-mouse Ig. Following washes, a label, such as I^{125} -ADMP, is added to each well and followed by incubation. Positive wells are subsequently detected by autoradiography. Positive clones can be grown in bulk culture and the supernatant subsequently purified utilizing a Protein A column.

Other types of "antibodies" may be produced using the information provided herein in conjunction with the

state of knowledge in the art. Humanized antibodies that are capable of specifically binding ADMPs are also encompassed by the instant invention.

Once isolated and purified, the antibodies against ADMPs can be used to detect the presence of ADMPs in a sample using established assay protocols. The antibodies of the invention can also be used therapeutically to bind to an ADMP and inhibit its activity in vivo.

The purified ADMPs according to the invention will facilitate the discovery of inhibitors of aggrecanases, and thus, inhibitors of cartilage aggrecan degradation. The use of a purified ADMP polypeptide in the screening of potential inhibitors thereof is important and can virtually eliminate the possibility of interfering reactions with contaminants. Such a screening assay for detecting the aggrecanase-inhibiting activity of a molecule would typically involve mixing the potential inhibitor molecule with an appropriate substrate, incubating an ADMP that is at least substantially purified with the mixture, and determining the extent of substrate cleavage. While various appropriate substrates may be designed for use in the assay, preferably the native aggrecan monomer or a peptidyl substrate which encompasses the E^{374-374A} cleavage site within the interglobular domain of the aggrecan core protein.

Alternatively, monitoring cleavage at aggrecanase-sensitive sites within the C-terminus of the aggrecan core protein, including E^{1545-1546G}, E^{1714-1715G}, E^{1819-1820A}, E^{1919-1920L} (numbering based on human aggrecan-core protein), can be used for detecting aggrecanase-inhibiting activity of a molecule by employing appropriate peptidyl substrates or the native aggrecan monomer and neoepitope antibodies.

In addition, ADMP polypeptides can be used for structure-based design of aggrecanase-inhibitors. Such structure-based design is also known as "rational drug design." The ADMP polypeptides can be three-dimensionally analyzed by X-ray crystallography, nuclear

magnetic resonance or homology modeling. The use of ADMP structural information in molecular modeling software systems to assist in the inhibitor design and inhibitor-ADMP interaction is also encompassed by the instant invention. Such computer-assisted modeling and drug design can utilize information such as chemical conformational analysis, electrostatic potential of the molecules, and protein folding. A particular method of the invention comprises analyzing the three-dimensional structure of ADMPs for likely binding sites of substrates, synthesizing a new molecule incorporating a predictive reactive site, and assaying the new molecule as described above.

The following examples provide an illustration of embodiments of the invention and should not be construed to limit the scope of the invention which is set forth in the appended claims. In the following examples, all methods described are conventional unless otherwise specified.

EXAMPLE 1

Purification of ADMP-1

This example describes a method for purifying ADMP-1. ADMP-1 was isolated and purified from the conditioned media of stimulated bovine nasal cartilage. Thirty liters of conditioned media, from approximately 1000 grams of bovine nasal cartilage, was generated by stimulating with interleukin-1 β (IL-1). In order to accumulate ADMPs in culture media, cartilage matrix was first degraded and depleted of endogenous aggrecan by stimulation with 500 ng/ml human recombinant IL-1 for 6 days with media changes every 2 days. Cartilage was then stimulated for an additional 10 days with 500 ng/ml IL-1 to generate accumulation of soluble, active ADMPs in the media. By replacing the media and restimulating the cartilage with IL-1 every other day during the accumulation phase, approximately 5 times more aggrecanase activity was generated than by allowing accumulation of ADMPs in conditioned media without media change. Media, containing the ADMPs, was frozen at -70°C for subsequent purification.

All purification steps were performed at 4° C unless otherwise specified. Five liters of frozen conditioned media was thawed overnight and supplemented with 1 μ M leupeptin, 1 μ M pepstatin, 1 mM PMSF (PMSF is phenylmethylsulfonyl-fluoride), and 0.05% Brij-35. This was clarified by passage through a 1.2 micron Gelman Capsule filter, and loaded onto a 20 x 10 cm Macro S support column at a flow rate of 40 ml/min. The column was washed with Buffer A (Buffer A contains 50 mM HEPES (HEPES is N-(2-hydroxyethyl) piperazine-N'-(2-ethanesulfonic acid), pH 7.5, 10 mM CaCl₂, 100 mM NaCl, 0.05% (v/v) Brij-35) until the absorbance at 280 nm returned to the pre-load baseline value. ADMPs were eluted from the column with 750 ml of buffer A containing 1.0 M NaCl.

ADMPs were detected at this point by their ability to cleave purified bovine aggrecan monomers isolated from

bovine nasal cartilage by the following procedure: Aggrecan was extracted from the cartilage by stirring at 4°C for 48 hours in 10 volumes of 4M guanidine-HCl in 0.05 M sodium acetate, pH 5.8, containing the protease inhibitors, 0.01M EDTA (EDTA is ethylenediaminetetraacetic acid), 0.1M 6-aminohexanoic acid, 2mM PMSF and 0.05M benzamidine HCl. Aggrecan monomers were isolated by equilibrium density gradient centrifugation in cesium chloride [Hascall, V.C. and Sajdera, S.W. (1969) *J. Biol. Chem.* 244, 2384-2396.] and the bottom of this gradient ($d > 1.54$ g/ml) containing the aggrecan monomers, was dialyzed at 4°C against water and lyophilized.

These aggrecan monomers (500 nM) were incubated at 37°C for at least 4 hr with ADMPs eluted from the Macro S support column in a final volume of 200 ul in Buffer B (Buffer B contains 50 mM Tris, pH 7.6, containing 0.1 M NaCl and 10 mM CaCl₂), quenched with 20 mM EDTA and analyzed for aggrecan fragments produced exclusively by cleavage at the Glu³⁷³-Ala³⁷⁴ bond within the aggrecan core protein using the monoclonal antibody, BC-3 (Hughes, C.E., et al., *Biochem. J.* 306:799-804, 1995). This antibody recognizes aggrecan fragments with the N-terminal sequence A³⁷⁴RGSVIL..., generated upon cleavage by ADMPs. The BC-3 antibody recognizes this neoepitope only when it is the N-terminus and not when it is present internally within aggrecan fragments or within the intact aggrecan core protein. Other proteases produced by cartilage in response to stimulation of chondrocytes do not cleave at the Glu³⁷³-Ala³⁷⁴ site, therefore only products produced upon cleavage by ADMPs are detected.

Removal of the glycosaminoglycan (GAG) side chains from aggrecan is necessary for the BC-3 antibody to recognize the ARGSVIL epitope of the core protein. Therefore, to remove GAGs from the aggrecan, samples were enzymatically deglycosylated with chondroitinase ABC (#EC4.2.2.4; Seikagaku Co., Kogyo, Japan) 0.1 units/10 ug GAG in Buffer D (Buffer D contains 50 mM sodium

acetate, pH 6.5, 100 mM NaCl) for 2 hr at 37°C and then with keratanase (#EC3.2.1.103; Seikagaku Co., Kogyo, Japan) (0.1 units/10 ug GAG) and keratanase II (Seikagaku Co., Kogyo, Japan) (0.002 units/10 ug GAG) in Buffer D for 2 hr at 37°C.

After digestion, the samples were precipitated with 5 volumes of acetone and reconstituted in an appropriate volume of SDS-PAGE sample buffer, loaded on 4-12% gradient gels and then separated by SDS-PAGE under non reducing conditions, transferred overnight to nitrocellulose and immunolocalized with 1:1000 dilution of the monoclonal antibody BC-3. Subsequently, membranes were incubated with goat anti-mouse IgG alkaline phosphatase conjugate and aggrecan catabolites visualized by incubation with the appropriate substrate (#S3721; Promega Western blot alkaline phosphatase system) for 10-30 min to achieve optimal color development. BC-3-reactive aggrecan fragments were then quantified by scanning densitometry.

The material that eluted from the Macro S support column with 0.1 M NaCl had about a 20-fold higher specific activity than the starting material. The eluted material was supplemented with Compound A (N3-methyl-(3R)-2-[(2S)-2-[(1R)-2-(hydroxyamino)-1-methyl-2-oxoethyl]-4-methylpentanoyl]hexahydro-3-pyridazinecarboxamide), a hydroxamic acid-based broad spectrum inhibitor of matrix metalloproteinases that is ineffective as an inhibitor of ADMP activity. The sample was then loaded onto a 10 x 7.5 cm gelatin-agarose column at a flow rate of 0.5 ml/min. Compound A was added to prevent degradation of the gelatin column by matrix metalloproteinases present during this purification step. Material passing through this column contained the ADMP activity and was collected and concentrated 6 to 7-fold using an Amicon Diaflo pressure concentrator fitted with a YM-30 membrane.

ADMP activity is inhibited by both tissue inhibitor of metalloproteinases-1 (TIMP-1) and by a number of hydroxamic acid-based inhibitors of matrix metalloproteinases.

Therefore TIMP-1 and a hydroxamate inhibitor of aggrecanase activity were used to further affinity purify ADMP. Compound A was included during the affinity purification to prevent matrix metalloproteinases present in the material from binding to the TIMP-1 or to the hydroxamate inhibitor of ADMP.

The concentrated material from the gelatin-agarose column containing the ADMP activity was incubated with 1 uM bovine TIMP-1 in the presence of 1 uM Compound A in for at least 30 minutes to allow ADMP to bind to the TIMP-1. The TIMP-1 was subsequently complexed by incubating with a TIMP-1 monoclonal antibody at a 1:5000 dilution for at least 30 minutes. The TIMP-1-antibody complex was then applied to a 10 ml protein A column. The column was washed 3X with Buffer E (Buffer E contains 10 mM Tris, pH 7.5, 250 mM NaCl, 0.025% Tween20) and the protein was eluted from the column with 100 mM glycine/HCl, pH 2.5.

An ADMP-inhibitor hydroxamate affinity resin was produced in the following manner. POROS 20-NH perfusive chromatographic media (Perseptive Biosystems), a highly crosslinked poly(styrene divinylbenzene) polymer, was mixed overnight with four equivalents of Fmoc-beta-alanine, four equivalents HBTU (O-Benzotriazole-N,N,N',N'-tetramethyl-uronium hexafluorophosphate), and eight equivalents DIEA (N,N diisopropylethylamine) in DMF (dimethylformamide). The resin was rinsed several times with DMF, then the Fmoc group was removed with 20% piperidine/DMF. This coupling/deprotecting scheme was repeated three more times, resulting in four beta-alanine residues coupled to the resin. After the final piperidine deprotection, 1.5 equivalents of a t-butyl protected hydroxamic acid capable of inhibiting ADMP activity was coupled with 1.5 equivalents HBTU and 3 equivalents DIEA in DMF. After rinsing with DMF and CH₂Cl₂, the t-butyl group was removed by mixing overnight with TFA (trifluoroacetic acid), leaving the hydroxamic

acid. The resin was thoroughly rinsed with CH_2Cl_2 and dried under vacuum.

The eluate from the protein A column was neutralized with 1 M Tris base to pH 7.5 and then incubated for 2 hours with the aggrecanase-inhibitor hydroxamate resin at a ratio of 1 mg resin for every 1 ml of eluate. Following the incubation, the resin was spun down and washed at 4°C with Buffer E, three times, 10 minutes each wash. Bound ADMP was eluted from the resin by mixing with approximately 0.5-1.0 ml 4 M GuHCl for 30 minutes at room temperature. The eluate, containing the ADMP activity, was dialyzed against Buffer F (Buffer F comprises 50mM Tris, pH 7.6, 100mM NaCl, 5mM CaCl_2) for 48 hours at 4°C. A portion of the eluate was run on a SDS-PAGE gel (10-20%), and silver staining revealed one predominant protein, ADMP-1, that ran as a doublet between the 64 and 92 kDa markers on the gel at approximately 67 kDa.

Incubation of the GuHCl-eluted ADMP-1 with isolated bovine aggrecan produced a pattern of BC-3-reactive fragments similar to that produced upon cleavage of cartilage aggrecan by endogenous aggrecanase in bovine nasal cartilage stimulated with IL-1.

Binding of this 67 kDa ADMP-1 doublet to the aggrecanase-inhibitor hydroxamate resin was blocked by inclusion during affinity purification of 10 μM Compound B ((2S,11S, 12S)-12-isobutyl-2-[(methylamino)carbonyl]-11[(hydroxylamino)carbonyl]-8,13-dioxo-1,7-diazacyclotridecane), a potent ADMP inhibitor. Thus, the binding of the 67 kDa ADMP-1 to the affinity resin was not affected by the presence of Compound A (a potent inhibitor of matrix metalloproteinases that is inactive in inhibiting ADMPs), but was blocked by Compound B (a potent ADMP inhibitor).

A region of the SDS-PAGE gel containing the 67 kDa ADMP-1 protein was excised along with a control region of the same gel which did not contain any detectable protein. Gel slices were incubated with 1% Triton X-100

for 1 hour at room temperature to remove the SDS for the gel. The gel was then crushed in 1 ml final volume of Buffer G (Buffer G contains 50 mM Tris, pH 7.5, 100 mM NaCl, 10 mM CaCl₂) and dialyzed in 10 kDa cutoff dialysis membrane for 48 hours against 12 liters of Buffer B at 4°C. ADMP activity was then determined by incubation with aggrecan substrate and monitoring products using the BC-3 antibody. The region of the gel containing the 67 kDa ADMP-1 protein exhibited activity while that of the control portion of the gel did not.

Generation of peptides and sequencing

The two protein bands running as a doublet at approximately 67 kDa and identified with ADMP activity by elution from SDS gels, were immobilized on PVDF and subjected to N-terminal amino acid sequence analysis as follows. The N-terminal sequence analyses were performed using the Hewlett-Packard G1005A N-terminal sequencer [Miller, C.G. (1994) Methods: A Companion to Methods in Enzymology 6, 315-333]. These analyses of PVDF-blotted protein were carried out using a modification of the Hewlett-Packard Routine 3.0 Sequencing Methods [Hewlett-Packard technical note TNF95-1: Routine 3.0 Sequencing Methods] in combination with the Hewlett-Packard PTH 4.M HPLC Method [Hewlett-Packard technical note TN95-6: PTH 4.M HPLC Method]. The modifications of the sequencing methods involved: (1) the replacement of the lower SAX sample column with a lower RP adapter column and (2) the substitution of a mixture of LC-A buffer and water (LC-A:water; 2:1) for solvent S4 at the step where the PTH-amino acids were resuspended for injection into the HPLC. All sequencer reagents were from Hewlett-Packard and the HPLC-grade water used to dilute the LC-A as described above was from Aldrich Chemical Company.

Both bands of the 67 kDa doublet represented the same protein and the sequence of the first 27 residues were determined to be:

SEQ ID NO:4 FASLSRFVETLVVADDKMAAFHGAGLK

Two internal fragments, a 7-mer and a 11-mer, from a tryptic digestion of the 67 kDa ADMP-1 protein were sequenced and have the following sequences:

SEQ ID NO:5 YTGVA PR

SEQ ID NO:6 ALGYYYVLDP

EXAMPLE 2

Purification of ADMP-2

This example describes a method for purifying a second member of the ADMP family, ADMP-2. The concentrated material from the gelatin-agarose column containing the ADMP activity prepared as detailed in Example 1 was loaded onto a 3 x 20 cm phenyl-sepharose column that had been equilibrated with buffer A containing 10% (w/v) ammonium sulfate and without the Brij-35 detergent. The column was then washed with 150 ml of equilibration buffer. Proteins were eluted with a 400 ml gradient from 10% to 0% ammonium sulfate in equilibration buffer at a flow rate of 2 ml/min. At the end of the gradient phase, proteins were further eluted from the column with an additional 200 ml of 0% ammonium sulfate equilibration buffer. Fractions were collected throughout the loading, washing and elution phases and analyzed for ADMP activity. Fractions containing ADMP activity were pooled.

Material that eluted from the phenyl sepharose column with ammonium sulfate had about 87-fold higher specific activity than the starting material. These pooled fractions were loaded onto a 4 x 60cm CM column that had been equilibrated with buffer H (buffer H

contains 50 mM HEPES, .1 M NaCl, pH 7.5). The column was washed with 150 ml of buffer H, and then proteins were eluted from the column with a 300 ml gradient from 0.1 to 1.0 M NaCl in buffer H at a flow rate of 1.5 ml/min. Fractions were collected throughout the loading, washing, and elution phases and analyzed for ADMP activity. Fractions with high enzymatic activity were pooled. While still containing contaminating proteins, the enzyme at this stage was of adequate purity for use in high throughput screens designed to find small molecular weight inhibitors of ADMP activity.

The pooled fractions from the CM column containing ADMP activity had about a 500 fold higher specific activity than the starting material. This material was concentrated approximately 10-fold with an Amicon Diaflo pressure concentrator fitted with a YM-30 membrane. Samples of 2 ml of the concentrate were then applied to a 2 x 200 cm Sephacryl S-200 column equilibrated with Buffer H. The column was eluted isocratically in the same buffer at a flow rate of 0.2 ml/min. Fractions of 4 ml were collected throughout the elution and analyzed for ADMP activity. Pooled fractions containing ADMP activity had about a 3500-fold higher specific activity than the starting material.

The ADMP eluted from the sizing column was concentrated as described above to a final volume of 1 ml and injected onto a (30 mm x 4.6 mm) C₄ alkylsilane-derivatized silica column and eluted with a linear gradient from 0 to 50% (v/v) acetonitrile in 0.1% aqueous TFA, at a flow rate of 1 ml/min. Fractions were collected throughout the elution phase, and immediately diluted 10-fold with Buffer A. As long as the acetonitrile concentration was diluted quickly, good recovery of enzymatic activity was observed from this column. Pooled fractions containing ADMP activity had about 100,000-fold higher specific activity than the starting material.

Analysis of the HPLC purified ADMP by SDS-PAGE with silver staining demonstrated the presence of multiple protein bands ranging in apparent molecular weights from approximately 90 to 30 kDa. Prominent bands were observed in the range of 65 to 40 kDa. To identify which protein band(s) corresponded to the ADMP activity, two samples of the HPLC purified ADMP were electrophoretically fractionated in adjacent wells of a 10% Tris-glycine polyacrylamide gel under non-reducing conditions. One lane was stained with silver. The other lane was cut horizontally into 22 approximately equal volume slices, each representing a different molecular weight range. The individual slices were crushed and soaked in 100 μ l of 20 mM Tris, 10 mM CaCl_2 100 mM NaCl, 2.5% Triton-X100. The samples were incubated in this buffer at 4° C overnight to renature the enzyme and elute it from the gel slice into the supernatant solution. The resulting solutions were tested for ADMP activity. ADMP activity was associated with four protein bands, centered at ca. 64, 62, 54 and 50 kDa.

Generation of peptides and sequencing

The four protein bands identified with ADMP activity by elution from SDS gels were immobilized on PVDF and subjected to N-terminal amino acid sequence analysis. The 64 kDa protein yielded 41 cycles of sequence having the following composition:

SEQ ID NO:16 SISRARQVEL LLVADASMAR MYGRGLQHYL
LTLASIANKLYF

The 62, 54, and 50 kDa protein bands yielded the same N-terminal amino acid sequence as found in the 64 kDa band.

Based on these results it appeared that the 64, 62, 54 and 50 kDa protein bands displaying ADMP activity represent different forms of the same ADMP-2 protein. The difference in molecular weights of these four bands is most likely the result of differential processing of

the four forms (e.g., differential glycosylation or C-terminal proteolytic processing).

EXAMPLE 3

Cloning of Human ADMP-1

This example describes a procedure for isolating a DNA sequence encoding human ADMP-1. N-terminal sequence obtained from the 67-kDa protein was shown to be 62% identical to the murine ADAMTS-1 protein (K. Kuno, et al. J. Biol. Chem. 272:556-562, 1997), suggesting that the 67 kDa protein was a member of the ADAM family of proteins. As defined herein and in the reference above, the name "ADAMTS-1" is an abbreviation for A disintegrin and metalloproteinase with thrombospondin motifs. Upon subsequent internal peptide sequencing, SEQ ID NO:6 was shown to be 50% identical to ADAMTS-1 and 91% (10/11) identical to sequences encoded by a murine EST 474985 (Accession number AA041973). Based on the high degree of identity between the latter internal peptide sequence and sequences encoded by EST474985, we proceeded to clone the human sequences representing the human homologue of murine EST 474985.

Two approaches were utilized in the cloning of the human homologue of murine EST 474985. In the first, the cDNA from which EST474985 was obtained was sequenced in its entirety, with the resulting sequence being used to search the EST data base for human ESTs with significant homology. In the second approach, human sequences were amplified using PCR primers designed from murine EST 474985 (SEQ ID NO:7 and SEQ ID NO:8). Both approaches provided us with sequence from the human homologue. A human EST (Accession number D45652) was identified that contained sequences having significant homology to the 1.7 kb murine cDNA from which EST 474985 was derived. The level of identity was 80% overall at the nucleotide level and contained sequences from the non-coding 3' untranslated region. Utilizing the second approach, we

were able to obtain a 190 bp amplicon from human heart cDNA. Subsequent DNA sequence analysis indicated that the human PCR product was 89% identical at the nucleotide level to the murine EST, with the deduced peptide sequences being 96% identical.

We were successful at obtaining additional coding sequences utilizing PCR primers designed from the human EST and the 190 bp PCR product. A 2-kb clone was obtained using sense and anti-sense primers designed from the 190-bp PCR product (SEQ ID NO:9) and the human 3' EST (SEQ ID NO:10), respectively. Utilizing antisense primers designed from the 190-bp PCR product (SEQ ID NO:11 and SEQ ID NO:12), we were able to clone a 2.2-kb 5' RACE (rapid amplification of cDNA ends) product. Six clones from the 5' and 3' PCRs were sequenced in order to obtain a consensus sequence for the cDNA. A total 4.2-kb of sequence has been assembled (SEQ ID NO:1). The assembled cDNA contains a 2511-bp open reading frame encoding 837 amino acids (SEQ ID NO:2) with multiple in-frame stop codons being present upstream of the start methionine. The cDNA encodes sequences present in the N-terminal peptide sequence (SEQ ID NO:4), with the cDNA containing all 27 residues seen in the N-terminal sequence of the 67 kDa ADMP-1 protein. The N-terminus of the isolated 67 kDa ADMP-1 begins with amino acid 213 of SEQ ID NO:2, indicating that this protein lacks the propeptide domain of the molecule. The cDNA also encodes sequences present in the internal peptide sequences (SEQ ID NO:5 and SEQ ID NO:6), with sequences encoded by the cDNA being identical in six of seven positions for the first peptide and ten of eleven positions for the second peptide. The deduced protein sequence has homology to the previously reported ADAMTS-1 protein. Like ADAMTS-1, the ADMP-1 cDNA contains a propeptide domain, metalloproteinase domain and disintegrin-like domain. A noteworthy difference between the ADAMTS-1 and sequences encoded by the ADMP-1 cDNA is the presence of a single thrombospondin-domain in the deduced aggrecan degrading

metallo protease, in contrast to the three thrombospondin domains seen in ADAMTS-1.

EXAMPLE 4

Cloning of ADMP-2

This example describes a procedure for isolating a DNA sequence encoding human ADMP-2. A closely related family member was identified by using the deduced ADMP-1 peptide sequence to search the EST data base. Sequences encoded by murine EST 569515 were shown to be 67% identical to sequences encoded by the ADMP-1 cDNA sequence. Further sequence analysis of the murine cDNA from which the EST 569515 was obtained indicated that it encoded sequences that were 95% (39/41) identical and 100% similar to the N-terminal sequence of the purified 50/64 kD aggrecanase, ADMP-2. These data indicated that the murine cDNA encodes the ADMP-2. PCR primers (SEQ ID NO: 17 and SEQ ID NO: 18) were designed from the murine sequence and were used to amplify a product from human heart cDNA, with the resulting product being 92% identical (150/163) at the nucleotide and 100% identical at the amino acid level to the murine sequences. Sequences present in the human amplicon were used to design a PCR primer for use in 3' RACE (SEQ ID NO: 19). However, only a partial 3' RACE clones was obtained using this approach. In order to obtain additional sequences, a human liver cDNA library was screened by PCR as described (D.I. Israel (1993) Nuc. Acids Res. 21, 2627-2631) using PCR primers designed from the human amplicon and the partial 3' RACE clone (SEQ ID NO: 19 and SEQ ID NO: 20). Two cDNA clones were obtained from the liver library, with each clone being approximately 5.5 kb in size.

Sequence analysis of the cDNA clones indicated that both cDNAs contain a 2793 bp open reading frame (SEQ ID NO: 14) encoding a 930 amino acid protein (SEQ ID NO: 15). The deduced protein sequence contains sequences that are 97.5% (40/41) identical and 100% similar to the bovine N-terminal peptide sequence of ADMP-2. The predicted protein encodes

an ADMP family member closely related to human ADMP-1 and to the murine ADAMTS-1. All three proteins contain metalloproteinase-domains, disintegrin-like domains and thrombospondin motifs. These three family members were found to have a variable number of thrombospondin-submotifs. Murine ADAMTS-1 has been shown to contain two thrombospondin -submotifs (Kuno et al (1997) J. Biol. Chem 272, 556-562), while the ADMP-2 cDNA encodes one thrombospondin-submotif and the ADMP-1 lacks the thrombospondin-submotifs altogether. Overall, the pro-domains of the three proteins are the least conserved, with the percent identity ranging from 15% for the ADMP-1 and ADMP-2 to 33% for mADAMTS-1 and ADMP-1. Greatest conservation was seen in the catalytic domains with the percent identity ranging from 48% (ADMP-1 and ADMP-2) to 62% (mADAMTS-1 and ADMP-1).

EXAMPLE 5

Preparation of Antibodies Against ADMP-1

This example describes a method for generating antibodies against ADMP-1. A peptide based on the N-terminus of the purified 67 kDa ADMP-1 was synthesized with the following sequence:

SEQ ID NO:13 CASLSRFVETLVVADDK

The peptide was linked to the carrier protein, keyhole limpet hemocyanin, and then subsequently used for immunization of a sheep. The coupled peptide antigen was suspended in PBS (phosphate-buffered saline) at 1 mg/ml with an equal volume of complete Freund's adjuvant. The material was mixed until it formed an emulsion, and then the material was injected at 6-8 subcutaneous sites. A total of 150-200 ug of coupled peptide was injected into the animal. The sheep was boosted every two weeks (for a

total of five times) and a production bleed was collected at each time point. The affinity of the antibody was tested both in an ELISA and Western assay using the above antigen peptide conjugated to BSA. The polyclonal antiserum was positive for recognizing the BSA coupled peptide both in the ELISA and Western assays. The polyclonal sera was affinity purified over an antigen peptide (CASLSRFVETLVVADDK) column to capture the high affinity IgG antibodies and remove the low affinity antibodies.

Example 6

Preparation of Antibodies Against ADMP-2

This example describes a method for generating antibodies against ADMP-2. A peptide based on the N-terminus of the purified 50/64 kDa ADMP-2 was synthesized with the following sequence:

SEQ ID NO: 21: SISRARQVELLAhx-C-amide

Aminohexanoic acid (Ahx) was added to lengthen the peptide. The peptide was linked to the carrier protein, keyhole limpet hemocyanin, and then subsequently used for immunization of two rabbits. Five immunizations and bleedings yielded 200-250 ml of serum. The affinity of the antibody was compared with preimmune sera using an enzyme linked immunosorbent assay (ELISA) with BSA-coupled peptide on the solid phase. For all sera, results were expressed as the reciprocal of the serum dilution that resulted in an OD₄₀₅ of 0.2 by detection with alkaline phosphatase-anti-rabbit IgG conjugate and NPP dye. The polyclonal antiserum was positive for recognizing the BSA coupled peptide in the ELISA and 90% of the aggrecanase activity in a 41-mer peptide-based enzymatic assay could be immunoprecipitated by the antibody. The polyclonal sera was affinity purified over an antigen peptide column to capture the high affinity

IgG antibodies and remove the low affinity antibodies.
The resulting affinity-purified antibody works well for
recognizing ADMP-2 bands in Western analysis.

What is claimed:

1. An isolated ADMP, cleaving the Glu³⁷³-Ala³⁷⁴ peptide bond of an aggrecan core protein.
2. An isolated ADMP, wherein said ADMP has the property of cleaving the Glu³⁷³-Ala³⁷⁴ peptide bond of the aggrecan core protein and wherein said ADMP consists of the following domains from the N-terminus to the C-terminus: 1) a propeptide domain containing a furin cleavage site, followed by 2) a metalloprotease domain, followed by 3) a disintegrin-like domain, and 4) a thrombospondin homologous domain.
3. A non-glycosylated ADMP according to claim 1 or 2 in isolated form.
4. An ADMP according to claim 1 or 2, having a molecular weight between about 35kD and about 120kD, as measured by SDS-PAGE.
5. An ADMP according to claim 1 or 2, having a molecular weight between about 45kD and about 100kD, as measured by SDS-PAGE.
6. An ADMP according to claim 1 or 2 having a molecular weight between about 50kD and about 98kD, as measured by SDS-PAGE.
7. An isolated nucleic acid selected from the group comprising:
 - (a) a coding region of a gene encoding a native mammalian ADMP of claim 1;
 - (b) a nucleic acid that is at least 80% identical to the nucleic acid of (a) and that encodes an ADMP;or

- (c) a nucleic acid which is degenerate as a result of the genetic code to nucleic acid defined in (a) or (b) and which encodes an ADMP.
8. An isolated nucleic acid according to claim 7, wherein the ADMP is a human ADMP.
 9. An isolated nucleic acid of claim 7 or 8 which encodes an active ADMP variant or ADMP derivative.
 10. An ADMP according to claim 1 or 2, that has a molecular weight of about 98kD for the zymogen as measured by the SDS-PAGE.
 11. An ADMP according to claim 1 or 2, that has a molecular weight of about 67kD as measured by SDS-PAGE for the mature form of the protease lacking the propeptide domain.
 12. An isolated nucleic acid selected from the group comprising:
 - (a) a coding region of a gene encoding native mammalian ADMP of claim 10;
 - (b) a cDNA comprising nucleotides 406-2919 of SEQ ID NO:1;
 - (c) a nucleic acid that is at least 80% identical to the nucleic acid of (a) or (b) and that encodes an ADMP; or
 - (d) a nucleic acid which is degenerate as a result of the genetic code to nucleic acid defined in (a), (b) or (c) and which encodes an ADMP.
 13. An isolated nucleic acid according to claim 12, wherein the ADMP is a human ADMP.
 14. An ADMP according to claim 1 or 2 that has a molecular weight of about 93kD for the zymogen as measured by SDS-PAGE.

15. An ADMP according to claim 1 or 2, that has a molecular weight of about 50kD, 54kDa, 62kDa or 64kD as measured by SDS-PAGE for the mature protease lacking the propeptide domain.
16. A composition comprising an ADMP as in any of claims 1 or 2.
17. An isolated nucleic acid selected from the group comprising:
 - (a) a coding region of a gene encoding native mammalian ADMP having a molecular weight of about 93kD as measured by SDS-PAGE;
 - (b) a cDNA comprising nucleotides 121-2910 of SEQ ID NO:14;
 - (c) a nucleic acid that is at least 80% identical to the nucleic acid of (a) or (b) and that encodes an ADMP and
 - (d) a nucleic acid which is degenerate as a result of the genetic code to nucleic acid defined in (a), (b) or (c) and which encodes an ADMP.
18. An isolated nucleic acid according to claim 17, wherein the ADMP is a human ADMP.
19. An antibody that binds to an ADMP according to claim 1.
20. An antibody according to claim 19, wherein the antibody is a monoclonal antibody.
21. An expression vector that directs the expression of a nucleic acid sequence as in any of claims 7, 8, 12, 13, 17 or 18.
22. A host cell transfected or transformed with an

expression vector that directs the expression of a nucleic acid sequence as in any of claims 7,8,12,13,17 or 18.

23. A method for producing an ADMP, comprising:
 - (a) culturing a host cell under conditions promoting expression; and
 - b) recovering the ADMP from the culture medium.
24. ADMPs produced according to claim 23.
25. ADMP-1 comprising the sequence of amino acids 1-837 of SEQ.ID NO:2.
26. ADMP-2 comprising the sequence of amino acids 1-930 of SEQ.ID NO:15.
27. A method of identifying cell lines, cells or tissues that produce ADMPs using a nucleic acid probe which hybridizes with a native ADMP nucleic acid sequence according to any of claims 7,8,12,13,17 or 18 to detect ADMP message in biological samples.
28. The use of an ADMP according to any of claims 1 or 2 or 25 or 26 for three-dimensional structural analysis and computer-aided drug design of ADMP inhibitors.
29. A method of inhibiting the cleavage of aggrecan in a mammal comprising administering to such mammal an effective amount of a compound that inhibits the proteolytic activity of an ADMP as in any of claims 1 or 2 or 25 or 26.
30. A method for treating a mammal having a disease characterized by an overproduction or an up-regulated production of an ADMP comprising administering to the mammal a pharmaceutical

composition comprising an amount of a compound that effectively inhibits the proteolytic activity of an ADMP as in any of claims 1 or 2 or 25 or 26.

31. A method of inhibiting the cleavage of aggrecan in a mammal comprising administering to such mammal an effective amount of a compound that inhibits the isolated ADMP comprising the sequence of amino acids 1-837 of SEQ ID NO:2.
32. A method for treating a mammal having a disease characterized by an overproduction or an up-regulated production of an ADMP, comprising administering to the mammal a composition comprising an amount of a compound that effectively inhibits the ADMP activity of an enzyme comprising the sequence of amino acids 1-837 of SEQ ID NO:2.
33. A method of inhibiting the cleavage of aggrecan in a mammal comprising administering to such mammal an effective amount of a compound that inhibits the isolated ADMP comprising the sequence of amino acids 1-930 of SEQ ID NO:15.
34. A method of inhibiting the cleavage of aggrecan in a mammal comprising administering to such mammal an effective amount of a compound that inhibits the ADMP activity of an enzyme wherein the protein used is ADAMTS-1.
35. A method for treating a mammal having a disease characterized by an overproduction or an upregulated production of an ADMP, comprising administering to the mammal a composition comprising an amount of a compound that effectively inhibits the ADMP activity of an enzyme comprising the sequence of amino acids 1-930 of SEQ ID NO:15.

36. A method for treating a mammal having a disease characterized by an overproduction or an up-regulated production of an ADMP, comprising administering to the mammal a composition comprising an amount of a compound that effectively inhibits the ADMP activity of an enzyme wherein the protein used is ADAMTS-1.
37. An assay for ADMP activity and inhibitors of ADMPs, which comprises:
- (a) generating soluble ADMPs, by stimulation of human or animal tissue or cells;
 - (b) detecting ADMP enzymatic activity by using the soluble ADMPs generated in (a) or any ADMP-containing sample, incubating with an aggrecan substrate in absence or presence of an inhibitor, and monitoring production of aggrecan fragments produced by specific cleavage at an ADMP-susceptible site using a neoepitope antibody to the new N-terminus or C-terminus generated by ADMP-mediated cleavage;
 - (c) evaluating inhibition of ADMP activity by comparing the amount of product produced in the presence versus absence of compound.
38. An assay according to claim 37 wherein the tissue or cell source of ADMPs is cartilage or chondrocytes.
39. An assay according to claim 37 or 38 wherein the aggrecan substrate is aggrecan isolated from human or animal tissue.
40. An assay according to claim 37 wherein the

aggrecan substrate is a recombinant aggrecan molecule or recombinant portion of the aggrecan molecule containing an aggrecanase-sensitive cleavage site.

41. An assay according to claim 40 wherein the recombinant portion of the aggrecan molecule contains the E³⁷³⁻³⁷⁴A bond.
42. An assay according to claim 40 wherein the recombinant aggrecan fragment contains the E¹⁵⁴⁵⁻¹⁵⁴⁶G bond.
43. An assay according to claim 40 wherein the portion of the aggrecan molecule contains the E¹⁷¹⁴⁻¹⁷¹⁵G bond.
44. An assay according to claim 40 wherein the recombinant portion of the aggrecan molecule contains the E¹⁸¹⁹⁻¹⁸²⁰A bond.
45. An assay according to claim 40 wherein the recombinant portion of the aggrecan molecule contains the E¹⁹¹⁹⁻¹⁹²⁰L bond.
46. An assay according to claims 37, 38, 40-45 wherein fragments produced by specific cleavage at the the E³⁷³⁻³⁷⁴A bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.
47. An assay according to claims 37, 38, 40-45 wherein fragments produced by specific cleavage at the the E¹⁵⁴⁵⁻¹⁵⁴⁶G bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.

48. An assay according to claims 37, 38, 40-45 wherein fragments produced by specific cleavage at the the E1714-1715G bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.
49. An assay according to claim 37, 38, 40-45 wherein fragments produced by specific cleavage at the the E1819-1820A bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond. .
50. An assay according to claim 37, 38, 40-45 wherein fragments produced by specific cleavage at the the E1919-1920L bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.
51. An assay according to claim 37, 38, 40-45 wherein the ARGSV N-terminus is detected using the monoclonal antibody, BC-3.
52. A method for assaying compounds for activity against an ADMP comprising:
- providing an ADMP and an ADMP substrate;
- contacting said ADMP with a candidate inhibitory compound in the presence of said ADMP; and
- measuring the inhibition of the ADMP activity.
53. A method for assaying compounds for activity against an ADAMTS-1 comprising:
- providing an ADAMTS-1 and an ADAMTS-1 substrate;

contacting said ADAMTS-1 with a candidate inhibitory compound in the presence of said ADAMTS-1; and

measuring the inhibition of the ADAMTS-1 activity.

54. A method for diagnosing arthritic diseases in a mammal using an antibody according to claim 19 or 20 to detect ADMPs in biological samples.
55. An method for detecting ADMP-generated products in biological samples which comprises detecting the new N-terminus or new C-terminus on fragments produced by specific cleavage at an ADMP-sensitive site using a neoepitope antibody to the N-terminus or C-terminus on the ADMP-generated fragment.
56. An assay according to claim 55 wherein fragments produced by specific cleavage at the the E373-374A bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.
57. An assay according to claim 55 wherein fragments produced by specific cleavage at the the E1545-1546G bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.
58. An assay according to claim 55 wherein fragments produced by specific cleavage at the E1714-1715G bond are monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.
59. An assay according to claim 55 wherein fragments produced by specific cleavage at the E1819-1820A bond are monitored using a neoepitope antibody to the new

C-terminus or new N-terminus on fragments generated by cleavage at this bond.

60. An assay according to claim 55 wherein produced by specific cleavage at the E¹⁹¹⁹-L¹⁹²⁰ bond is monitored using a neoepitope antibody to the new C-terminus or new N-terminus on fragments generated by cleavage at this bond.
61. A method for diagnosing arthritic diseases in a mammal by monitoring specific ADMP-generated aggrecan fragments in biological samples according to any of claims 55-60.
61. A method for diagnosing arthritic diseases in a mammal using a nucleic acid probe which hybridizes with a native ADMP nucleic acid sequence according to any of claims 7,8,12,13,17 or 18 to detect ADMP message in biological samples.
62. A method for diagnosing a disease in a mammal characterized by an overproduction or an up-regulated production of an ADMP using an antibody to detect ADMPs in biological samples.
63. A method for diagnosing a disease in a mammal characterized by an overproduction or an up-regulated production of an ADMP by monitoring specific ADMP-generated fragments in biological samples according to any of claims 55-60.
64. A method for diagnosing a disease in a mammal characterized by an overproduction or an up-regulated production of an ADMP using a nucleic acid probe which hybridizes with a native ADMP nucleic acid sequence according to any of claims 7,8,12,13,17 or 18 to detect ADMP message in biological samples.

ADMP-1 Sequences

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 4192 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURES:

- (A) NAME/KEY: CDS
- (B) LOCATION: 406..2919

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```

ACA GAC ACA TAT GCA CGA GAG AGA CAG AGG AGG AAA GAG ACA GAG ACA 48
AAG GCA CAG CGG AAG AAG GCA GAG ACA GGG CAG GCA CAG AAG CGG CCC 96
AGA CAG AGT CCT ACA GAG GGA GAG GCC AGA GAA GCT GCA GAA GAC ACA 144
GGC AGG GAG AGA CAA AGA TCC AGG AAA GGA GGG CTC AGG AGG AGA GTT 192
TGG AGA AGC CAG ACC CCT GGG CAC CTC TCC CAA GCC CAA GGA CTA AGT 240
TTT CTC CAT TTC CTT TAA CGG TCC TCA GCC CTT CTG AAA ACT TTG CCT 288
CTG ACC TTG GCA GGA GTC CAA GCC CCC AGG CTA CAG AGA GGA GCT TTC 336
CAA AGC TAG GGT GTG GAG GAC TTG GTG CCC TAG ACG GCC TCA GTC CCT 384
CCC AGC TGC AGT ACC AGT GCC ATG TCC CAG ACA GGC TCG CAT CCC GGG 432
                               Met Ser Gln Thr Gly Ser His Pro Gly
AGG GGC TTG GCA GGG CGC TGG CTG TGG GGA GCC CAA CCC TGC CTC CTG 480
Arg Gly Leu Ala Gly Arg Trp Leu Trp Gly Ala Gln Pro Cys Leu Leu
CTC CCC ATT GTG CCG CTC TCC TGG CTG GTG TGG CTG CTT CTG CTA CTG 528
Leu Pro Ile Val Pro Leu Ser Trp Leu Val Trp Leu Leu Leu Leu Leu
CTG GCC TCT CTC CTG CCC TCA GCC CGG CTG GCC AGC CCC CTC CCC CGG 576
Leu Ala Ser Leu Leu Pro Ser Ala Arg Leu Ala Ser Pro Leu Pro Arg
GAG GAG GAG ATC GTG TTT CCA GAG AAG CTC AAC GGC AGC GTC CTG CCT 624
Glu Glu Glu Ile Val Phe Pro Glu Lys Leu Asn Gly Ser Val Leu Pro
GGC TCG GGC GCC CCT GCC AGG CTG TTG TGC CGC TTG CAG GCC TTT GGG 672
Gly Ser Gly Ala Pro Ala Arg Leu Leu Cys Arg Leu Gln Ala Phe Gly
GAG ACG CTG CTA CTA GAG CTG GAG CAG GAC TCC GGT GTG CAG GTC GAG 720
Glu Thr Leu Leu Leu Glu Leu Glu Gln Asp Ser Gly Val Gln Val Glu
GGG CTG ACA GTG CAG TAC CTG GGC CAG GCG CCT GAG CTG CTG GGT GGA 768
Gly Leu Thr Val Gln Tyr Leu Gly Gln Ala Pro Glu Leu Leu Gly Gly
GCA GAG CCT GGC ACC TAC CTG ACT GGC ACC ATC AAT GGA GAT CCG GAG 816
Ala Glu Pro Gly Thr Tyr Leu Thr Gly Thr Ile Asn Gly Asp Pro Glu

```

TCG GTG GCA TCT CTG CAC TGG GAT GGG GGA GCC CTG TTA GGC GTG TTA 864
Ser Val Ala Ser Leu His Trp Asp Gly Gly Ala Leu Leu Gly Val Leu

CAA TAT CGG GGG GCT GAA CTC CAC CTC CAG CCC CTG GAG GGA GGC ACC 912
Gln Tyr Arg Gly Ala Glu Leu His Leu Gln Pro Leu Glu Gly Gly Thr

CCT AAC TCT GCT GGG GGA CCT GGG GCT CAC ATC CTA CGC CGG AAG AGT 960
Pro Asn Ser Ala Gly Gly Pro Gly Ala His Ile Leu Arg Arg Lys Ser

CCT GCC AGC GGT CAA GGT CCC ATG TGC AAC GTC AAG GCT CCT CTT GGA 1008
Pro Ala Ser Gly Gln Gly Pro Met Cys Asn Val Lys Ala Pro Leu Gly

AGC CCC AGC CCC AGA CCC CGA AGA GCC AAG CGC TTT GCT TCA CTG AGT 1056
Ser Pro Ser Pro Arg Pro Arg Arg Ala Lys Arg Phe Ala Ser Leu Ser

AGA TTT GTG GAG ACA CTG GTG GTG GCA GAT GAC AAG ATG GCC GCA TTC 1104
Arg Phe Val Glu Thr Leu Val Val Ala Asp Asp Lys Met Ala Ala Phe

CAC GGT GCG GGG CTA AAG CGC TAC CTG CTA ACA GTG ATG GCA GCA GCA 1152
His Gly Ala Gly Leu Lys Arg Tyr Leu Leu Thr Val Met Ala Ala Ala

GCC AAG GCC TTC AAG CAC CCA AGC ATC CGC AAT CCT GTC AGC TTG GTG 1200
Ala Lys Ala Phe Lys His Pro Ser Ile Arg Asn Pro Val Ser Leu Val

GTG ACT CGG CTA GTG ATC CTG GGG TCA GGC GAG GAG GGG CCC CAA GTG 1248
Val Thr Arg Leu Val Ile Leu Gly Ser Gly Glu Glu Gly Pro Gln Val

GGG CCC AGT GCT GCC CAG ACC CTG CGC AGC TTC TGT GCC TGG CAG CGG 1296
Gly Pro Ser Ala Ala Gln Thr Leu Arg Ser Phe Cys Ala Trp Gln Arg

GGC CTC AAC ACC CCT GAG GAC TCG GAC CCT GAC CAC TTT GAC ACA GCC 1344
Gly Leu Asn Thr Pro Glu Asp Ser Asp Pro Asp His Phe Asp Thr Ala

ATT CTG TTT ACC CGT CAG GAC CTG TGT GGA GTC TCC ACT TGC GAC ACG 1392
Ile Leu Phe Thr Arg Gln Asp Leu Cys Gly Val Ser Thr Cys Asp Thr

CTG GGT ATG GCT GAT GTG GGC ACC GTC TGT GAC CCG GCT CGG AGC TGT 1440
Leu Gly Met Ala Asp Val Gly Thr Val Cys Asp Pro Ala Arg Ser Cys

GCC ATT GTG GAG GAT GAT GGG CTC CAG TCA GCC TTC ACT GCT GCT CAT 1488
Ala Ile Val Glu Asp Asp Gly Leu Gln Ser Ala Phe Thr Ala Ala His

GAA CTG GGT CAT GTC TTC AAC ATG CTC CAT GAC AAC TCC AAG CCA TGC 1536
Glu Leu Gly His Val Phe Asn Met Leu His Asp Asn Ser Lys Pro Cys

ATC AGT TTG AAT GGG CCT TTG AGC ACC TCT CGC CAT GTC ATG GCC CCT 1584
Ile Ser Leu Asn Gly Pro Leu Ser Thr Ser Arg His Val Met Ala Pro

GTG ATG GCT CAT GTG GAT CCT GAG GAG CCC TGG TCC CCC TGC AGT GCC 1632
Val Met Ala His Val Asp Pro Glu Glu Pro Trp Ser Pro Cys Ser Ala

CGC TTC ATC ACT GAC TTC CTG GAC AAT GGC TAT GGG CAC TGT CTC TTA 1680
Arg Phe Ile Thr Asp Phe Leu Asp Asn Gly Tyr Gly His Cys Leu Leu

GAC AAA CCA GAG GCT CCA TTG CAT CTG CCT GTG ACT TTC CCT GGC AAG 1728
Asp Lys Pro Glu Ala Pro Leu His Leu Pro Val Thr Phe Pro Gly Lys

GAC TAT GAT GCT GAC CGC CAG TGC CAG CTG ACC TTC GGG CCC GAC TCA 1776
Asp Tyr Asp Ala Asp Arg Gln Cys Gln Leu Thr Phe Gly Pro Asp Ser

CGC CAT TGT CCA CAG CTG CCG CCG CCC TGT GCT GCC CTC TGG TGC TCT 1824
Arg His Cys Pro Gln Leu Pro Pro Pro Cys Ala Ala Leu Trp Cys Ser

GGC CAC CTC AAT GGC CAT GCC ATG TGC CAG ACC AAA CAC TCG CCC TGG	1872
Gly His Leu Asn Gly His Ala Met Cys Gln Thr Lys His Ser Pro Trp	
GCC GAT GGC ACA CCC TGC GGG CCC GCA CAG GCC TGC ATG GGT GGT CGC	1920
Ala Asp Gly Thr Pro Cys Gly Pro Ala Gln Ala Cys Met Gly Gly Arg	
TGC CTC CAC ATG GAC CAG CTC CAG GAC TTC AAT ATT CCA CAG GCT GGT	1968
Cys Leu His Met Asp Gln Leu Gln Asp Phe Asn Ile Pro Gln Ala Gly	
GGC TGG GGT CCT TGG GGA CCA TGG GGT GAC TGC TCT CGG ACC TGT GGG	2016
Gly Trp Gly Pro Trp Gly Pro Trp Gly Asp Cys Ser Arg Thr Cys Gly	
GGT GGT GTC CAG TTC TCC TCC CGA GAC TGC ACG AGG CCT GTC CCC CGG	2064
Gly Gly Val Gln Phe Ser Ser Arg Asp Cys Thr Arg Pro Val Pro Arg	
AAT GGT GGC AAG TAC TGT GAG GGC CGC CGT ACC CGC TTC CGC TCC TGC	2112
Asn Gly Gly Lys Tyr Cys Glu Gly Arg Arg Thr Arg Phe Arg Ser Cys	
AAC ACT GAG GAC TGC CCA ACT GGC TCA GCC CTG ACC TTC CGC GAG GAG	2160
Asn Thr Glu Asp Cys Pro Thr Gly Ser Ala Leu Thr Phe Arg Glu Glu	
CAG TGT GCT GCC TAC AAC CAC CGC ACC GAC CTC TTC AAG AGC TTC CCA	2208
Gln Cys Ala Ala Tyr Asn His Arg Thr Asp Leu Phe Lys Ser Phe Pro	
GGG CCC ATG GAC TGG GTT CCT CGC TAC ACA GGC GTG GCC CCC CAG GAC	2256
Gly Pro Met Asp Trp Val Pro Arg Tyr Thr Gly Val Ala Pro Gln Asp	
CAG TGC AAA CTC ACC TGC CAG GCC CGG GCA CTG GGC TAC TAC TAT GTG	2304
Gln Cys Lys Leu Thr Cys Gln Ala Arg Ala Leu Gly Tyr Tyr Tyr Val	
CTG GAG CCA CGG GTG GTA GAT GGG ACC CCC TGT TCC CCG GAC AGC TCC	2352
Leu Glu Pro Arg Val Val Asp Gly Thr Pro Cys Ser Pro Asp Ser Ser	
TCG GTC TGT GTC CAG GGC CGA TGC ATC CAT GCT GGC TGT GAT CGC ATC	2400
Ser Val Cys Val Gln Gly Arg Cys Ile His Ala Gly Cys Asp Arg Ile	
ATT GGC TCC AAG AAG AAG TTT GAC AAG TGC ATG GTG TGC GGA GGG GAC	2448
Ile Gly Ser Lys Lys Lys Phe Asp Lys Cys Met Val Cys Gly Gly Asp	
GGT TCT GGT TGC AGC AAG CAG TCA GGC TCC TTC AGG AAA TTC AGG TAC	2496
Gly Ser Gly Cys Ser Lys Gln Ser Gly Ser Phe Arg Lys Phe Arg Tyr	
GGA TAC AAC AAT GTG GTC ACT ATC CCC GCG GGG GCC ACC CAC ATT CTT	2544
Gly Tyr Asn Asn Val Val Thr Ile Pro Ala Gly Ala Thr His Ile Leu	
GTC CGG CAG CAG GGA AAC CCT GGC CAC CGG AGC ATC TAC TTG GCC CTG	2592
Val Arg Gln Gln Gly Asn Pro Gly His Arg Ser Ile Tyr Leu Ala Leu	
AAG CTG CCA GAT GGC TCC TAT GCC CTC AAT GGT GAA TAC ACG CTG ATG	2640
Lys Leu Pro Asp Gly Ser Tyr Ala Leu Asn Gly Glu Tyr Thr Leu Met	
CCC TCC CCC ACA GAT GTG GTA CTG CCT GGG GCA GTC AGC TTG CGC TAC	2688
Pro Ser Pro Thr Asp Val Val Leu Pro Gly Ala Val Ser Leu Arg Tyr	
AGC GGG GCC ACT GCA GCC TCA GAG ACA CTG TCA GGC CAT GGG CCA CTG	2736
Ser Gly Ala Thr Ala Ala Ser Glu Thr Leu Ser Gly His Gly Pro Leu	
GCC CAG CCT TTG ACA CTG CAA GTC CTA GTG GCT GGC AAC CCC CAG GAC	2784
Ala Gln Pro Leu Thr Leu Gln Val Leu Val Ala Gly Asn Pro Gln Asp	
ACA CGC CTC CGA TAC AGC TTC TTC GTG CCC CGG CCG ACC CCT TCA ACG	2832
Thr Arg Leu Arg Tyr Ser Phe Phe Val Pro Arg Pro Thr Pro Ser Thr	
CCA CGC CCC ACT CCC CAG GAC TGG CTG CAC CGA AGA GCA CAG ATT CTG	2880

Pro Arg Pro Thr Pro Gln Asp Trp Leu His Arg Arg Ala Gln Ile Leu
 GAG ATC CTT CGG CGG CGC CCC TGG GCG GGC AGG AAA TAA CCT CAC TAT 2928
 Glu Ile Leu Arg Arg Arg Pro Trp Ala Gly Arg Lys End
 CCC GGC TGC CCT TTC TGG GCA CCG GGG CCT CGG ACT TAG CTG GGA GAA 2976
 AGA GAG AGC TTC TGT TGC TGC CTC ATG CTA AGA CTC AGT GGG GAG GGG 3024
 CTG TGG GCG TGA GAC CTG CCC CTC CTC TCT GCC CTA ATG CGC AGG CTG 3072
 GCC CTG CCC TGG TTT CCT GCC CTG GGA GGC AGT GAT GGG TTA GTG GAT 3120
 GGA AGG GGC TGA CAG ACA GCC CTC CAT CTA AAC TGC CCC CTC TGC CCT 3168
 GCG GGT CAC AGG AGG GAG GGG GAA GGC AGG GAG GGC CTG GGC CCC AGT 3216
 TGT ATT TAT TTA GTA TTT ATT CAC TTT TAT TTA GCA CCA GGG AAG GGG 3264
 ACA AGG ACT AGG GTC CTG GGG AAC CTG ACC CCT GAC CCC TCA TAG CCC 3312
 TCA CCC TGG GGC TAG GAA ATC CAG GGT GGT GGT GAT AGG TAT AAG TGG 3360
 TGT GTG TAT GCG TGT GTG TGT GTG TGT GAA AAT GTG TGT GTG CTT ATG 3408
 TAT GAG GTA CAA CCT GTT CTG CTT TCC TCT TCC TGA ATT TTA TTT TTT 3456
 GGG AAA AGA AAA GTC AAG GGT AGG GTG GGC CTT CAG GGA GTG AGG GAT 3504
 TAT CCT TTT TTT TTT CTT TCT TTC TTT CTT TTT TTT TTT GAG ACA GAA 3552
 TCT CGC TCT GTC GCC CAG GCT GGA GTG CAA TGG CAC AAT CTC GGC TCA 3600
 CTG CAT CCT CCG CCT CCC GGG TTC AAG TGA TTC TCA TGC CTC AGC CTC 3648
 CTG AGT AGC TGG GAT TAC AGG CTC CTG CCA CCA CGC CCG GCT AAT TTT 3696
 TGT TTT GTT TTG TTT GGA GAC AGA GTC TCG CTA TTG TCA CCA GGG CTG 3744
 GAA TGA TTT CAG CTC ACT GCA ACC TTC GCC ACC TGG GTT CCA GCA ATT 3792
 CTC CTG CCT CAG CCT CCC GAG TAG CTG AGA TTA TAG GCA CCT ACC ACC 3840
 ACG CCC GGC TAA TTT TTG TAT TTT TAG TAG AGA CGG GGT TTC ACC ATG 3888
 TTG GCC AGG CTG GTC TCG AAC TCC TGA CCT TAG GTG ATC CAC TCG CCT 3936
 TCA TCT CCC AAA GTG CTG GGA TTA CAG GCG TGA GCC ACC GTG CCT GGC 3984
 CAC GCC CAA CTA ATT TTT GTA TTT TTA GTA GAG ACA GGG TTT CAC CAT 4032
 GTT GGC CAG GCT GCT CTT GAA CTC CTG ACC TCA GGT AAT CGA CCT GCC 4080
 TCG GCC TCC CAA AGT GCT GGG ATT ACA GGT GTG AGC CAC CAC GCC CGG 4128
 TAC ATA TTT TTT AAA TTG AAT TCT ACT ATT TAT GTG ATC CTT TTG GAG 4176
 TCA GAC AGA TGT GGG T 4192

- (2) INFORMATION FOR SEQ ID NO:2:
 (i) SEQUENCE CHARACTERISTICS

(A) LENGTH: 837 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ser Gln Thr Gly Ser His Pro Gly Arg Gly Leu Ala Gly Arg Trp 16
 Leu Trp Gly Ala Gln Pro Cys Leu Leu Leu Pro Ile Val Pro Leu Ser 32
 Trp Leu Val Trp Leu Leu Leu Leu Leu Leu Ala Ser Leu Leu Pro Ser 48
 Ala Arg Leu Ala Ser Pro Leu Pro Arg Glu Glu Glu Ile Val Phe Pro 64
 Glu Lys Leu Asn Gly Ser Val Leu Pro Gly Ser Gly Ala Pro Ala Arg 80
 Leu Leu Cys Arg Leu Gln Ala Phe Gly Glu Thr Leu Leu Leu Glu Leu 96
 Glu Gln Asp Ser Gly Val Gln Val Glu Gly Leu Thr Val Gln Tyr Leu 112
 Gly Gln Ala Pro Glu Leu Leu Gly Gly Ala Glu Pro Gly Thr Tyr Leu 128
 Thr Gly Thr Ile Asn Gly Asp Pro Glu Ser Val Ala Ser Leu His Trp 144
 Asp Gly Gly Ala Leu Leu Gly Val Leu Gln Tyr Arg Gly Ala Glu Leu 160
 His Leu Gln Pro Leu Glu Gly Gly Thr Pro Asn Ser Ala Gly Gly Pro 176
 Gly Ala His Ile Leu Arg Arg Lys Ser Pro Ala Ser Gly Gln Gly Pro 192
 Met Cys Asn Val Lys Ala Pro Leu Gly Ser Pro Ser Pro Arg Pro Arg 208
 Arg Ala Lys Arg Phe Ala Ser Leu Ser Arg Phe Val Glu Thr Leu Val 224
 Val Ala Asp Asp Lys Met Ala Ala Phe His Gly Ala Gly Leu Lys Arg 240
 Tyr Leu Leu Thr Val Met Ala Ala Ala Ala Lys Ala Phe Lys His Pro 256
 Ser Ile Arg Asn Pro Val Ser Leu Val Val Thr Arg Leu Val Ile Leu 272
 Gly Ser Gly Glu Glu Gly Pro Gln Val Gly Pro Ser Ala Ala Gln Thr 288
 Leu Arg Ser Phe Cys Ala Trp Gln Arg Gly Leu Asn Thr Pro Glu Asp 304
 Ser Asp Pro Asp His Phe Asp Thr Ala Ile Leu Phe Thr Arg Gln Asp 320
 Leu Cys Gly Val Ser Thr Cys Asp Thr Leu Gly Met Ala Asp Val Gly 336
 Thr Val Cys Asp Pro Ala Arg Ser Cys Ala Ile Val Glu Asp Asp Gly 352
 Leu Gln Ser Ala Phe Thr Ala Ala His Glu Leu Gly His Val Phe Asn 368
 Met Leu His Asp Asn Ser Lys Pro Cys Ile Ser Leu Asn Gly Pro Leu 384
 Ser Thr Ser Arg His Val Met Ala Pro Val Met Ala His Val Asp Pro 400
 Glu Glu Pro Trp Ser Pro Cys Ser Ala Arg Phe Ile Thr Asp Phe Leu 416
 Asp Asn Gly Tyr Gly His Cys Leu Leu Asp Lys Pro Glu Ala Pro Leu 432
 His Leu Pro Val Thr Phe Pro Gly Lys Asp Tyr Asp Ala Asp Arg Gln 448

Cys Gln Leu Thr Phe Gly Pro Asp Ser Arg His Cys Pro Gln Leu Pro 464
 Pro Pro Cys Ala Ala Leu Trp Cys Ser Gly His Leu Asn Gly His Ala 480
 Met Cys Gln Thr Lys His Ser Pro Trp Ala Asp Gly Thr Pro Cys Gly 496
 Pro Ala Gln Ala Cys Met Gly Gly Arg Cys Leu His Met Asp Gln Leu 512
 Gln Asp Phe Asn Ile Pro Gln Ala Gly Gly Trp Gly Pro Trp Gly Pro 528
 Trp Gly Asp Cys Ser Arg Thr Cys Gly Gly Gly Val Gln Phe Ser Ser 544
 Arg Asp Cys Thr Arg Pro Val Pro Arg Asn Gly Gly Lys Tyr Cys Glu 560
 Gly Arg Arg Thr Arg Phe Arg Ser Cys Asn Thr Glu Asp Cys Pro Thr 576
 Gly Ser Ala Leu Thr Phe Arg Glu Glu Gln Cys Ala Ala Tyr Asn His 592
 Arg Thr Asp Leu Phe Lys Ser Phe Pro Gly Pro Met Asp Trp Val Pro 608
 Arg Tyr Thr Gly Val Ala Pro Gln Asp Gln Cys Lys Leu Thr Cys Gln 624
 Ala Arg Ala Leu Gly Tyr Tyr Tyr Val Leu Glu Pro Arg Val Val Asp 640
 Gly Thr Pro Cys Ser Pro Asp Ser Ser Ser Val Cys Val Gln Gly Arg 656
 Cys Ile His Ala Gly Cys Asp Arg Ile Ile Gly Ser Lys Lys Lys Phe 672
 Asp Lys Cys Met Val Cys Gly Gly Asp Gly Ser Gly Cys Ser Lys Gln 688
 Ser Gly Ser Phe Arg Lys Phe Arg Tyr Gly Tyr Asn Asn Val Val Thr 704
 Ile Pro Ala Gly Ala Thr His Ile Leu Val Arg Gln Gln Gly Asn Pro 720
 Gly His Arg Ser Ile Tyr Leu Ala Leu Lys Leu Pro Asp Gly Ser Tyr 736
 Ala Leu Asn Gly Glu Tyr Thr Leu Met Pro Ser Pro Thr Asp Val Val 752
 Leu Pro Gly Ala Val Ser Leu Arg Tyr Ser Gly Ala Thr Ala Ala Ser 768
 Glu Thr Leu Ser Gly His Gly Pro Leu Ala Gln Pro Leu Thr Leu Gln 784
 Val Leu Val Ala Gly Asn Pro Gln Asp Thr Arg Leu Arg Tyr Ser Phe 800
 Phe Val Pro Arg Pro Thr Pro Ser Thr Pro Arg Pro Thr Pro Gln Asp 816
 Trp Leu His Arg Arg Ala Gln Ile Leu Glu Ile Leu Arg Arg Arg Pro 832
 Trp Ala Gly Arg Lys 837

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 27 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

FASLSRVETLVVADDKMAAFHGAGLK

- (2) INFORMATION FOR SEQ ID NO:5:
 (i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 7 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (ii) MOECULE TYPE: peptide
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

YTGVAAPR

- (2) INFORMATION FOR SEQ ID NO:6:
 (i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 27 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear
 (ii) MOECULE TYPE: peptide
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

ALGYYYVLDPR

- (2) INFORMATION FOR SEQ ID NO:7:
 (i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (iii) HYPOTHETICAL: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GGGGGTGGTGTCCAGTTCTCC

- (2) INFORMATION FOR SEQ ID NO:8:
 (i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (iii) HYPOTHETICAL: NO
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GGCCCTGGAAAGCTCTGAAGAG

- (2) INFORMATION FOR SEQ ID NO:9:
 (i) SEQUENCE CHARACTERISTICS
 (A) LENGTH: 23 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear
 (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCCCGGAATGGTGGCAAGTACTG

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ACCCACATCTGTCTGACTCCAAA

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCAGTTGGGCAGTCCTCAGTGTT

(2) INFORMATION FOR SEQ ID NO:12:

- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GGTCGGTGCGGTGGTTGTAGGC

(2) INFORMATION FOR SEQ ID NO:13:

- (i) SEQUENCE CHARACTERISTICS
 - (A) LENGTH: 17 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

(ii) MOECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CASLSRFVETLVVADDK

ADMP-2 Sequences

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 3250 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(ix) FEATURES:

- (A) NAME/KEY: CDS
- (B) LOCATION: 121..2913

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TGA CTC AAT CCT GCA AGC AAG TGT GTG TGT GTC CCC ATC CCC CGC CCC 48
 GTT AAC TTC ATA GCA AAT AAC AAA TAC CCA TAA AGT CCC AGT CGC GCA 96
 GCC CCT CCC CGC GGG CAG CGC ACT ATG CTG CTC GGG TGG GCG TCC CTG 144
 Met Leu Leu Gly Trp Ala Ser Leu
 CTG CTG TGC GCG TTC CGC CTG CCC CTG GCC GCG GTC GGC CCC GCC GCG 192
 Leu Leu Cys Ala Phe Arg Leu Pro Leu Ala Ala Val Gly Pro Ala Ala
 ACA CCT GCC CAG GAT AAA GCC GGG CAG CCT CCG ACT GCT GCA GCA GCC 240
 Thr Pro Ala Gln Asp Lys Ala Gly Gln Pro Pro Thr Ala Ala Ala Ala
 GCC CAG CCC CGC CGG CGG CAG GGG GAG GAG GTG CAG GAG CGA GCC GAG 288
 Ala Gln Pro Arg Arg Arg Gln Gly Glu Glu Val Gln Glu Arg Ala Glu
 CCT CCC GGC CAC CCG CAC CCC CTG GCG CAG CGG CGC AGG AGC AAG GGG 336
 Pro Pro Gly His Pro His Pro Leu Ala Gln Arg Arg Arg Ser Lys Gly
 CTG GTG CAG AAC ATC GAC CAA CTC TAC TCC GGC GGC GGC AAG GTG GGC 384
 Leu Val Gln Asn Ile Asp Gln Leu Tyr Ser Gly Gly Gly Lys Val Gly
 TAC CTC GTC TAC GCG GGC GGC CGG AGG TTC CTC TTG GAC CTG GAG CGA 432
 Tyr Leu Val Tyr Ala Gly Gly Arg Arg Phe Leu Leu Asp Leu Glu Arg
 GAT GGT TCG GTG GGC ATT GCT GGC TTC GTG CCC GCA GGA GGC GGG ACG 480
 Asp Gly Ser Val Gly Ile Ala Gly Phe Val Pro Ala Gly Gly Gly Thr
 AGT GCG CCC TGG CGC CAC CGG AGC CAC TGC TTC TAT CGG GGC ACA GTG 528
 Ser Ala Pro Trp Arg His Arg Ser His Cys Phe Tyr Arg Gly Thr Val
 GAC GCT AGT CCC CGC TCT CTG GCT GTC TTT GAC CTC TGT GGG GGT CTC 576
 Asp Ala Ser Pro Arg Ser Leu Ala Val Phe Asp Leu Cys Gly Gly Leu
 GAC GGC TTC TTC GCG GTC AAG CAC GCG CGC TAC ACC CTA AAG CCA CTG 624
 Asp Gly Phe Phe Ala Val Lys His Ala Arg Tyr Thr Leu Lys Pro Leu
 CTG CGC GGA CCC TGG GCG GAG GAA GAA AAG GGG CGC GTG TAC GGG GAT 672
 Leu Arg Gly Pro Trp Ala Glu Glu Glu Lys Gly Arg Val Tyr Gly Asp
 GGG TCC GCA CGG ATC CTG CAC GTC TAC ACC CGC GAG GGC TTC AGC TTC 720
 Gly Ser Ala Arg Ile Leu His Val Tyr Thr Arg Glu Gly Phe Ser Phe

GAG GCC CTG CCG CCG CGC GCC AGC TGC GAA ACC CCC GCG TCC ACA CCG 768
 Glu Ala Leu Pro Pro Arg Ala Ser Cys Glu Thr Pro Ala Ser Thr Pro
 GAG GCC CAC GAG CAT GCT CCG GCG CAC AGC AAC CCG AGC GGA CGC GCA 816
 Glu Ala His Glu His Ala Pro Ala His Ser Asn Pro Ser Gly Arg Ala
 GCA CTG GCC TCG CAG CTC TTG GAC CAG TCC GCT CTC TCG CCC GCT GGG 864
 Ala Leu Ala Ser Gln Leu Leu Asp Gln Ser Ala Leu Ser Pro Ala Gly
 GGC TCA GGA CCG CAG ACG TGG TGG CGG CGG CGG CGC CGC TCC ATC TCC 912
 Gly Ser Gly Pro Gln Thr Trp Trp Arg Arg Arg Arg Arg Ser Ile Ser
 CGG GCC CGC CAG GTG GAG CTG CTT CTG GTG GCT GAC GCG TCC ATG GCG 960
 Arg Ala Arg Gln Val Glu Leu Leu Leu Val Ala Asp Ala Ser Met Ala
 CGG TTG TAT GGC CGG GGC CTG CAG CAT TAC CTG CTG ACC CTG GCC TCC 1008
 Arg Leu Tyr Gly Arg Gly Leu Gln His Tyr Leu Leu Thr Leu Ala Ser
 ATC GCC AAT AGG CTG TAC AGC CAT GCT AGC ATC GAG AAC CAC ATC CGC 1056
 Ile Ala Asn Arg Leu Tyr Ser His Ala Ser Ile Glu Asn His Ile Arg
 CTG GCC GTG GTG AAG GTG GTG GTG CTA GGC GAC AAG GAC AAG AGC CTG 1104
 Leu Ala Val Val Lys Val Val Val Leu Gly Asp Lys Asp Lys Ser Leu
 GAA GTG AGC AAG AAC GCT GCC ACC ACA CTC AAG AAC TTT TGC AAG TGG 1152
 Glu Val Ser Lys Asn Ala Ala Thr Thr Leu Lys Asn Phe Cys Lys Trp
 CAG CAC CAA CAC AAC CAG CTG GGA GAT GAC CAT GAG GAG CAC TAC GAT 1200
 Gln His Gln His Asn Gln Leu Gly Asp Asp His Glu Glu His Tyr Asp
 GCA GCT ATC CTG TTT ACT CGG GAG GAT TTA TGT GGG CAT CAT TCA TGT 1248
 Ala Ala Ile Leu Phe Thr Arg Glu Asp Leu Cys Gly His His Ser Cys
 GAC ACC CTG GGA ATG GCA GAC GTT GGG ACC ATA TGT TCT CCA GAG CGC 1296
 Asp Thr Leu Gly Met Ala Asp Val Gly Thr Ile Cys Ser Pro Glu Arg
 AGC TGT GCT GTG ATT GAA GAC GAT GGC CTC CAC GCA GCC TTC ACT GTG 1344
 Ser Cys Ala Val Ile Glu Asp Asp Gly Leu His Ala Ala Phe Thr Val
 GCT CAC GAA ATC GGA CAT TTA CTT GGC CTC TCC CAT GAC GAT TCC AAA 1392
 Ala His Glu Ile Gly His Leu Leu Gly Leu Ser His Asp Asp Ser Lys
 TTC TGT GAA GAG ACC TTT GGT TCC ACA GAA GAT AAG CGC TTA ATG TCT 1440
 Phe Cys Glu Glu Thr Phe Gly Ser Thr Glu Asp Lys Arg Leu Met Ser
 TCC ATC CTT ACC AGC ATT GAT GCA TCT AAG CCC TGG TCC AAA TGC ACT 1488
 Ser Ile Leu Thr Ser Ile Asp Ala Ser Lys Pro Trp Ser Lys Cys Thr
 TCA GCC ACC ATC ACA GAA TTC CTG GAT GAT GGC CAT GGT AAC TGT TTG 1536
 Ser Ala Thr Ile Thr Glu Phe Leu Asp Asp Gly His Gly Asn Cys Leu
 CTG GAC CTA CCA CGA AAG CAG ATC CTG GGC CCC GAA GAA CTC CCA GGA 1584
 Leu Asp Leu Pro Arg Lys Gln Ile Leu Gly Pro Glu Glu Leu Pro Gly
 CAG ACC TAC GAT GCC ACC CAG CAG TGC AAC CTG ACA TTC GGG CCT GAG 1632
 Gln Thr Tyr Asp Ala Thr Gln Gln Cys Asn Leu Thr Phe Gly Pro Glu
 TAC TCC GTG TGT CCC GGC ATG GAT GTC TGT GCT CGC CTG TGG TGT GCT 1680
 Tyr Ser Val Cys Pro Gly Met Asp Val Cys Ala Arg Leu Trp Cys Ala
 GTG GTA CGC CAG GGC CAG ATG GTC TGT CTG ACC AAG AAG CTG CCT GCG 1728
 Val Val Arg Gln Gly Gln Met Val Cys Leu Thr Lys Lys Leu Pro Ala

GTG GAA GGG ACG CCT TGT GGA AAG GGG AGA ATC TGC CTG CAG GGC AAA 1776
 Val Glu Gly Thr Pro Cys Gly Lys Gly Arg Ile Cys Leu Gln Gly Lys
 TGT GTG GAC AAA ACC AAG AAA AAA TAT TAT TCA ACG TCA AGC CAT GGC 1824
 Cys Val Asp Lys Thr Lys Lys Lys Tyr Tyr Ser Thr Ser Ser His Gly
 AAC TGG GGA TCT TGG GGA TCC TGG GGC CAG TGT TCT CGC TCA TGT GGA 1872
 Asn Trp Gly Ser Trp Gly Ser Trp Gly Gln Cys Ser Arg Ser Cys Gly
 GGA GGA GTG CAG TTT GCC TAT CGT CAC TGT AAT AAC CCT GCT CCC AGA 1920
 Gly Gly Val Gln Phe Ala Tyr Arg His Cys Asn Asn Pro Ala Pro Arg
 AAC AAC GGA CGC TAC TGC ACA GGG AAG AGG GCC ATC TAC CGC TCC TG 1968C
 Asn Asn Gly Arg Tyr Cys Thr Gly Lys Arg Ala Ile Tyr Arg Ser Cys
 AGT CTC ATG CCC TGC CCA CCC AAT GGT AAA TCA TTT CGT CAT GAA CAG 2016
 Ser Leu Met Pro Cys Pro Pro Asn Gly Lys Ser Phe Arg His Glu Gln
 TGT GAG GCC AAA AAT GGC TAT CAG TCT GAT GCA AAA GGA GTC AAA ACT 2064
 Cys Glu Ala Lys Asn Gly Tyr Gln Ser Asp Ala Lys Gly Val Lys Thr
 TTT GTG GAA TGG GTT CCC AAA TAT GCA GGT GTC CTG CCA GCG GAT GTG 2112
 Phe Val Glu Trp Val Pro Lys Tyr Ala Gly Val Leu Pro Ala Asp Val
 TGC AAG CTG ACC TGC AGA GCC AAG GGC ACT GGC TAC TAT GTG GTA TTT 2160
 Cys Lys Leu Thr Cys Arg Ala Lys Gly Thr Gly Tyr Tyr Val Val Phe
 TCT CCA AAG GTG ACC GAT GGC ACT GAA TGT AGG CCG TAC AGT AAT TCC 2208
 Ser Pro Lys Val Thr Asp Gly Thr Glu Cys Arg Pro Tyr Ser Asn Ser
 GTC TGC GTC CGG GGG AAG TGT GTG AGA ACT GGC TGT GAC GGC ATC ATT 2209
 Val Cys Val Arg Gly Lys Cys Val Arg Thr Gly Cys Asp Gly Ile Ile
 GGC TCA AAG CTG CAG TAT GAC AAG TGC GGA GTA TGT GGA GGA GAC AAC 2304
 Gly Ser Lys Leu Gln Tyr Asp Lys Cys Gly Val Cys Gly Gly Asp Asn
 TCC AGC TGT ACA AAG ATT GTT GGA ACC TTT AAT AAG AAA AGT AAG GGT 2352
 Ser Ser Cys Thr Lys Ile Val Gly Thr Phe Asn Lys Lys Ser Lys Gly
 TAC ACT GAC GTG GTG AGG ATT CCT GAA GGG GCA ACC CAC ATA AAA GTT 2400
 Tyr Thr Asp Val Val Arg Ile Pro Glu Gly Ala Thr His Ile Lys Val
 CGA CAG TTC AAA GCC AAA GAC CAG ACT AGA TTC ACT GCC TAT TTA GCC 2448
 Arg Gln Phe Lys Ala Lys Asp Gln Thr Arg Phe Thr Ala Tyr Leu Ala
 CTG AAA AAG AAA AAC GGT GAG TAC CTT ATC AAT GGA AAG TAC ATG ATC 2496
 Leu Lys Lys Lys Asn Gly Glu Tyr Leu Ile Asn Gly Lys Tyr Met Ile
 TCC ACT TCA GAG ACT ATC ATT GAC ATC AAT GGA ACA GTC ATG AAC TAT 2544
 Ser Thr Ser Glu Thr Ile Ile Asp Ile Asn Gly Thr Val Met Asn Tyr
 AGC GGT TGG AGC CAC AGG GAT GAC TTC CTG CAT GGC ATG GGC TAC TCT 2592
 Ser Gly Trp Ser His Arg Asp Asp Phe Leu His Gly Met Gly Tyr Ser
 GCC ACG AAG GAA ATT CTA ATA GTG CAG ATT CTT GCA ACA GAC CCC ACT 2640
 Ala Thr Lys Glu Ile Leu Ile Val Gln Ile Leu Ala Thr Asp Pro Thr
 AAA CCA TTA GAT GTC CGT TAT AGC TTT TTT GTT CCC AAG AAG TCC ACT 2688
 Lys Pro Leu Asp Val Arg Tyr Ser Phe Phe Val Pro Lys Lys Ser Thr
 CCA AAA GTA AAC TCT GTC ACT AGT CAT GGC AGC AAT AAA GTG GGA TCA 2736
 Pro Lys Val Asn Ser Val Thr Ser His Gly Ser Asn Lys Val Gly Ser
 CAC ACT TCG CAG CCG CAG TGG GTC ACG GGC CCA TGG CTC GCC TGC TCT 2784

His Thr Ser Gln Pro Gln Trp Val Thr Gly Pro Trp Leu Ala Cys Ser
 AGG ACC TGT GAC ACA GGT TGG CAC ACC AGA ACG GTG CAG TGC CAG GAT 2832
 Arg Thr Cys Asp Thr Gly Trp His Thr Arg Thr Val Gln Cys Gln Asp
 GGA AAC CGG AAG TTA GCA AAA GGA TGT CCT CTC TCC CAA AGG CCT TCT 2880
 Gly Asn Arg Lys Leu Ala Lys Gly Cys Pro Leu Ser Gln Arg Pro Ser
 GCG TTT AAG CAA TGC TTG TTG AAG AAA TGT TAG CCT GTG GTT ATG ATC 2928
 Ala Phe Lys Gln Cys Leu Leu Lys Lys Cys
 TTA TGC ACA AAG ATA ACT GGA GGA TTC AGC ACC GAT GCA GTC GTG GTG 2976
 AAC AGG AGG TCT ACC TAA CGC ACA GAA AGT CAT GCT TCA GTG ACA TTG 3024
 TCA ACA GGA GTC CAA TTA TGG GCA GAA TCT GCT CTC TGT GAC CAA AAG 3072
 AGG ATG TGC ACT GCT TCA CGT GAC AGT GGT GAC CTT GCA ATA TAG AAA 3120
 AAC TTG GGA GTT ATT GAA CAT CCC CTG GGA TTA CAA GAA ACA CTG ATG 3168
 AAT GTT AAA TCA GGG GAC ATT TGA AGA TGG CAG AAC TGT CTC CCC CTT 3216
 GTC ACC TAC CTC TGA ATA GAA TGT CTT TAA TGG T 3250

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 930 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Leu Leu Gly Trp Ala Ser Leu Leu Leu Cys Ala Phe Arg Leu Pro 16
 Leu Ala Ala Val Gly Pro Ala Ala Thr Pro Ala Gln Asp Lys Ala Gly 32
 Gln Pro Pro Thr Ala Ala Ala Ala Gln Pro Arg Arg Arg Gln Gly 48
 Glu Glu Val Gln Glu Arg Ala Glu Pro Pro Gly His Pro His Pro Leu 64
 Ala Gln Arg Arg Arg Ser Lys Gly Leu Val Gln Asn Ile Asp Gln Leu 80
 Tyr Ser Gly Gly Gly Lys Val Gly Tyr Leu Val Tyr Ala Gly Gly Arg 96
 Arg Phe Leu Leu Asp Leu Glu Arg Asp Gly Ser Val Gly Ile Ala Gly 112
 Phe Val Pro Ala Gly Gly Gly Thr Ser Ala Pro Trp Arg His Arg Ser 128
 His Cys Phe Tyr Arg Gly Thr Val Asp Ala Ser Pro Arg Ser Leu Ala 144
 Val Phe Asp Leu Cys Gly Gly Leu Asp Gly Phe Phe Ala Val Lys His 160
 Ala Arg Tyr Thr Leu Lys Pro Leu Leu Arg Gly Pro Trp Ala Glu Glu 176
 Glu Lys Gly Arg Val Tyr Gly Asp Gly Ser Ala Arg Ile Leu His Val 192
 Tyr Thr Arg Glu Gly Phe Ser Phe Glu Ala Leu Pro Pro Arg Ala Ser 208
 Cys Glu Thr Pro Ala Ser Thr Pro Glu Ala His Glu His Ala Pro Ala 224

His Ser Asn Pro Ser Gly Arg Ala Ala Leu Ala Ser Gln Leu Leu Asp 240
 Gln Ser Ala Leu Ser Pro Ala Gly Gly Ser Gly Pro Gln Thr Trp Trp 256
 Arg Arg Arg Arg Arg Ser Ile Ser Arg Ala Arg Gln Val Glu Leu Leu 272
 Leu Val Ala Asp Ala Ser Met Ala Arg Leu Tyr Gly Arg Gly Leu Gln 288
 His Tyr Leu Leu Thr Leu Ala Ser Ile Ala Asn Arg Leu Tyr Ser His 304
 Ala Ser Ile Glu Asn His Ile Arg Leu Ala Val Val Lys Val Val Val 320
 Leu Gly Asp Lys Asp Lys Ser Leu Glu Val Ser Lys Asn Ala Ala Thr 336
 Thr Leu Lys Asn Phe Cys Lys Trp Gln His Gln His Asn Gln Leu Gly 352
 Asp Asp His Glu Glu His Tyr Asp Ala Ala Ile Leu Phe Thr Arg Glu 368
 Asp Leu Cys Gly His His Ser Cys Asp Thr Leu Gly Met Ala Asp Val 384
 Gly Thr Ile Cys Ser Pro Glu Arg Ser Cys Ala Val Ile Glu Asp Asp 400
 Gly Leu His Ala Ala Phe Thr Val Ala His Glu Ile Gly His Leu Leu 416
 Gly Leu Ser His Asp Asp Ser Lys Phe Cys Glu Glu Thr Phe Gly Ser 432
 Thr Glu Asp Lys Arg Leu Met Ser Ser Ile Leu Thr Ser Ile Asp Ala 448
 Ser Lys Pro Trp Ser Lys Cys Thr Ser Ala Thr Ile Thr Glu Phe Leu 464
 Asp Asp Gly His Gly Asn Cys Leu Leu Asp Leu Pro Arg Lys Gln Ile 480
 Leu Gly Pro Glu Glu Leu Pro Gly Gln Thr Tyr Asp Ala Thr Gln Gln 496
 Cys Asn Leu Thr Phe Gly Pro Glu Tyr Ser Val Cys Pro Gly Met Asp 512
 Val Cys Ala Arg Leu Trp Cys Ala Val Val Arg Gln Gly Gln Met Val 528
 Cys Leu Thr Lys Lys Leu Pro Ala Val Glu Gly Thr Pro Cys Gly Lys 544
 Gly Arg Ile Cys Leu Gln Gly Lys Cys Val Asp Lys Thr Lys Lys Lys 560
 Tyr Tyr Ser Thr Ser Ser His Gly Asn Trp Gly Ser Trp Gly Ser Trp 576
 Gly Gln Cys Ser Arg Ser Cys Gly Gly Gly Val Gln Phe Ala Tyr Arg 592
 His Cys Asn Asn Pro Ala Pro Arg Asn Asn Gly Arg Tyr Cys Thr Gly 608
 Lys Arg Ala Ile Tyr Arg Ser Cys Ser Leu Met Pro Cys Pro Pro Asn 624
 Gly Lys Ser Phe Arg His Glu Gln Cys Glu Ala Lys Asn Gly Tyr Gln 640
 Ser Asp Ala Lys Gly Val Lys Thr Phe Val Glu Trp Val Pro Lys Tyr 656
 Ala Gly Val Leu Pro Ala Asp Val Cys Lys Leu Thr Cys Arg Ala Lys 672
 Gly Thr Gly Tyr Tyr Val Val Phe Ser Pro Lys Val Thr Asp Gly Thr 688
 Glu Cys Arg Pro Tyr Ser Asn Ser Val Cys Val Arg Gly Lys Cys Val 704
 Arg Thr Gly Cys Asp Gly Ile Ile Gly Ser Lys Leu Gln Tyr Asp Lys 720
 Cys Gly Val Cys Gly Gly Asp Asn Ser Ser Cys Thr Lys Ile Val Gly 736

Thr Phe Asn Lys Lys Ser Lys Gly Tyr Thr Asp Val Val Arg Ile Pro 752
 Glu Gly Ala Thr His Ile Lys Val Arg Gln Phe Lys Ala Lys Asp Gln 768
 Thr Arg Phe Thr Ala Tyr Leu Ala Leu Lys Lys Lys Asn Gly Glu Tyr 784
 Leu Ile Asn Gly Lys Tyr Met Ile Ser Thr Ser Glu Thr Ile Ile Asp 800
 Ile Asn Gly Thr Val Met Asn Tyr Ser Gly Trp Ser His Arg Asp Asp 816
 Phe Leu His Gly Met Gly Tyr Ser Ala Thr Lys Glu Ile Leu Ile Val 832
 Gln Ile Leu Ala Thr Asp Pro Thr Lys Pro Leu Asp Val Arg Tyr Ser 848
 Phe Phe Val Pro Lys Lys Ser Thr Pro Lys Val Asn Ser Val Thr Ser 864
 His Gly Ser Asn Lys Val Gly Ser His Thr Ser Gln Pro Gln Trp Val 880
 Thr Gly Pro Trp Leu Ala Cys Ser Arg Thr Cys Asp Thr Gly Trp His 896
 Thr Arg Thr Val Gln Cys Gln Asp Gly Asn Arg Lys Leu Ala Lys Gly 912
 Cys Pro Leu Ser Gln Arg Pro Ser Ala Phe Lys Gln Cys Leu Leu Lys 928
 Lys Cys 930

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 42 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

SISRARQVELLLVADASMARMYGRGLQHYLLTLASIANKLYF

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGGCCACGACCCTCAAGAACTTT

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCATGGAGGCCATCATCTTCAATCA

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

GGGAGGATTTATGTGGGCATCA

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTGCATTGGACCAGGGCTTAGA

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 11 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

SISRARQVELLAhxC-amide

[The body of the document is mostly blank, with some faint, illegible markings and a few scattered characters.]